THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXVIII May, 1937 Number 5

RELATIONSHIPS BETWEEN FACTORS OBTAINED FROM CERTAIN ANALYSES

KARL J. HOLZINGER AND HARRY H. HARMAN

University of Chicago

1. INTRODUCTION

It is the purpose of this paper to note the characteristics of several of Thurstone's multiple factor analyses[1] together with our bi-factor solution,[2] and to show their relationship. Certain criteria will be given for judging factorial analyses and will be applied to various systems of factors. We shall illustrate the principles set forth in this paper by a short hypothetical problem in order to save space and to bring out the relationships of results more clearly.

From a geometric point of view, the tests may be considered as vectors with a common origin, and their intercorrelations as the scalar products[3] of these vectors or the cosines of the angles between these vectors normalized. The factor problem then becomes one of selecting a small number of reference vectors, or factors, to represent the tests appropriately. Before considering the various bases for selecting such reference axes,[4] we shall first distinguish two matrices which occur in factorial analysis.

One type of matrix, which we shall call a "factor structure" or more simply a "structure," is a table of correlations of the tests

[1] For a description of the methods of multiple factor analysis see *The Vectors of Mind*. Chicago: The University of Chicago Press, 1935.

[2] For a description of bi-factor analysis see *Preliminary Report on Spearman-Holzinger Unitary Trait Study*, No. 7. Prepared at the Statistical Laboratory, Department of Education, The University of Chicago, 1936.

[3] A term borrowed from vector analysis which means the product of the lengths of two vectors by the cosine of their angular separation.

[4] We use the terms reference vectors and reference axes interchangeably, the context indicating whether we are considering the directed line segment or the entire line.

with the factor vectors. The other consists of the coefficients in what we shall call a "factor pattern" or simply "pattern." A pattern may be defined as a set of equations expressing the tests as linear functions of the factors, where all variables and factors are in standardized form. It is necessary to show that a factor pattern can be formed from a factor structure before complete comparisons of different methods can be made.

A pattern may be written in the form,

$$x_j = f_{jA}F_A + f_{jB}F_B + f_{jC}F_C + \cdots + f_{jR}F_R, \qquad (j = 1, 2, 3, \cdots, n), \tag{1}$$

where x_j are the tests, $F_k (k = A, B, C, \cdots, R)$ are the factors, f_{jk} are coefficients, n is the number of tests, and R is the total number of factors. It is the matrix, (f_{jk}), obtained from the pattern that we distinguish from the structure.

Multiplying equations (1) by F_k, summing over the number of cases N, and dividing by N, we have,

$$\begin{aligned}
r_{x_jF_A} &= f_{jA} + f_{jB}r_{F_AF_B} + f_{jC}r_{F_AF_C} + \cdots + f_{jR}r_{F_AF_R}, \\
r_{x_jF_B} &= f_{jA}r_{F_BF_A} + f_{jB} + f_{jC}r_{F_BF_C} + \cdots + f_{jR}r_{F_BF_R}, \\
r_{x_jF_C} &= f_{jA}r_{F_CF_A} + f_{jB}r_{F_CF_B} + f_{jC} + \cdots + f_{jR}r_{F_CF_R}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
r_{x_jF_R} &= f_{jA}r_{F_RF_A} + f_{jB}r_{F_RF_B} + f_{jC}r_{F_RF_C} + \cdots + f_{jR}.
\end{aligned} \tag{2}$$

Equations (2) thus show the functional relationships between the elements of a structure $r_{x_jF_k}$ and the coefficients f_{jk} of a pattern.

If we consider equations (2) as n sets of R linearly independent[1] equations each, then we can solve[2] for the R coefficients f_{jk} in each of the n sets. Equations (2) thought of in this way involve the f_{jk}'s as unknowns and the elements of the structure $r_{x_jF_k}$ and the intercorrelations of factors $r_{F_kF_l}$ $(k, l = A, B, C, \cdots, R)$ as known numbers. The values f_{jk} may then be substituted in equations (1) to form a pattern. We have thus shown that we may construct a pattern from a known structure.

---

[1] We shall speak of a system of n equations $l_1 = k_1, \cdots, l_n = k_n$, where the k's are constants and $l_1, \ldots, l_n$ are linear forms in the unknowns $y_1, \ldots, y_n$, as *linearly dependent* if there exist numbers $c_1, \ldots, c_n$, not all zero, such that

$$c_1l_1 + \cdots + c_nl_n = 0.$$

If no such numbers $c_1, \ldots, c_n$ exist, we shall call the linear forms $l_1, \ldots, l_n$ (or the n equations) *linearly independent.*

[2] Dickson, L. E.: *Modern Algebraic Theories.* New York: Benj. H. Sanborn & Co., 1930, Theorem 16, p. 63.

It is interesting to note[1] that the values f_{jk} in the above solution turn out to be regression coefficients, so that equations (1) may be regarded as regression equations with the errors of estimate included. These errors are represented by the specific factors in the present case. Thus if there are M group factors, equations (1) may also be written in the form:

$$x_j = f_{jA}F_A + f_{jB}F_B + \cdots + f_{jM}F_M \pm f_{js}F_{js}, \qquad (1')$$

where F_{js} are the n specific factors, which may be split into two types if desired. The total number of factors R in equations (1) is equal to the total number of factors in the pattern exhibited in equations (1'), that is

$$R = M + n,$$

where M is the number of group factors and n is the number of specifics. We have placed a + or − in front of the f_{js} because the coefficients of the specific factors in a factor pattern are usually obtained, after all the other coefficients are known, in such a manner as to bring the variance of each test up to unity. This involves a square root and hence the plus or minus sign.

---

[1] If we denote the matrix of intercorrelations of factors, bordered by the correlations of any test x_j with the factors, by Δ, we have,

$$\Delta \equiv \begin{vmatrix} 1 & r_{x_jF_A} & r_{x_jF_B} & r_{x_jF_C} & \cdots & r_{x_jF_R} \\ r_{F_Ax_j} & 1 & r_{F_AF_B} & r_{F_AF_C} & \cdots & r_{F_AF_R} \\ r_{F_Bx_j} & r_{F_BF_A} & 1 & r_{F_BF_C} & \cdots & r_{F_BF_R} \\ r_{F_Cx_j} & r_{F_CF_A} & r_{F_CF_B} & 1 & \cdots & r_{F_CF_R} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{F_Rx_j} & r_{F_RF_A} & r_{F_RF_B} & r_{F_RF_C} & \cdots & 1 \end{vmatrix}.$$

Let Δ_{11} be the minor of the element in the first row and column and let Δ_{jk} be the cofactor of the element $r_{x_jF_k}$. Then the regression coefficients in the linear equation for x_j are given by

$$\beta_{jk\cdot ABC\cdots R} = \frac{-\Delta_{jk}}{\Delta_{11}}.$$

The subscripts on β are so arranged that the position of any regression coefficient is uniquely determined. Thus the primary subscript j indicates that x_j is the dependent variable; the next subscript k shows the factor F_k to which the coefficient is attached; and the remaining secondary subscripts following the decimal point merely show the number of variables involved.

Now the algebraic solution, by determinants, for the values f_{jk} from equations (2) yield,

$$f_{jk} = \beta_{jk\cdot ABC\cdots R}.$$

In all later factor patterns in this paper we shall write only that portion of the equations which involves common factors, and drop all specifics since they do not affect the succeeding work. Similarly in a matrix of coefficients of a pattern we shall enter only the coefficients of the common factors.

From equations (2) it is apparent that the elements of the structure, $r_{x_jF_k}$, are generally different from the coefficients f_{jk} of the pattern. In case the reference vectors F_k are mutually orthogonal, the correlations $r_{F_kF_l}(k \neq l)$ are zero. Equations (2) then reduce to

$$r_{x_jF_k} = f_{jk}. \tag{3}$$

It is thus apparent that only in the case of orthogonal factors are the elements of the structure identical with the corresponding coefficients of the pattern. This would indicate that the term "factor loading" should not be used in the case of oblique factors, or else should be defined to refer to either the elements of the structure or the pattern. There are examples in the literature where this term is used to apply to both types of elements with much confusion.

We are of the opinion that both structure and pattern should be produced in making a complete factor analysis involving oblique factors. The structure reveals the zero-order correlations of tests and factors, which are useful for the identification of factors, and for subsequent regression estimates of the latter. The pattern shows the composition of tests and factors in the form of regression equations. It is also necessary for the calculation of correlations between tests and of residuals to determine the adequacy of fit of the solution. Furthermore, patterns are necessary in order to compare factors obtained by different analyses.

The intercorrelations between tests can be obtained from a factor pattern by the following procedure: Multiply any equation (1′) of a test by the equation of another, sum over the number of cases N, and divide by N. For any two tests x_i, $x_j(i, j = 1, 2, 3, \cdots, n)$ we have,

$$\begin{aligned} r_{x_ix_j} = {} & f_{iA}f_{jA} + f_{iB}f_{jB} + f_{iC}f_{jC} + \cdots + f_{iM}f_{jM} \\ & + (f_{iA}f_{jB} + f_{jA}f_{iB})r_{F_AF_B} + (f_{iA}f_{jC} + f_{jA}f_{iC})r_{F_AF_C} + \cdots \\ & + (f_{iA}f_{jM} + f_{jA}f_{iM})r_{F_AF_M} + (f_{iB}f_{jC} + f_{jB}f_{iC})r_{F_BF_C} + \cdots \\ & + (f_{iB}f_{jM} + f_{jB}f_{iM})r_{F_BF_M} + \cdots + f_{iA}f_{jS}r_{F_AF_{jS}} \\ & + f_{jA}f_{iS}r_{F_AF_{iS}} + \cdots + f_{iM}f_{jS}r_{F_MF_{jS}} + f_{jM}f_{iS}r_{F_MF_{iS}} \\ & + f_{iS}f_{jS}r_{F_{iS}F_{jS}}. \end{aligned} \tag{4}$$

The axes of the specific factors are always at right angles to the common-factor space and also at right angles to each other. Hence, $r_{F_s F_{is}} = r_{F_s F_{is}} = r_{F_{is} F_{js}} = 0$. If the axes of the common factors are mutually orthogonal equation (4) simplifies still further. The correlations between factors $r_{F_s F_t}$ are then zero and everything below the first line of equation (4) vanishes.

2. CRITERIA FOR JUDGING FACTORIAL ANALYSES

We may now consider certain standards for judging the adequacy of a factorial analysis. The following criteria have been set forth explicitly or implicitly by several workers in this field.

a. Parsimony in Number of Factors.—This is an obvious desideratum of all factorial analyses. We wish to represent the highly complex configuration of original tests by a small number of correlated or uncorrelated factors.

b. Geometrical Fit.—If a group of tests show high correlations among themselves and low correlations with other tests, the vectors corresponding to this group will be close together and widely separated from other groups. A factor vector lying close to such a cluster would be one with good geometric fit.

c. Simplicity of Structure.—It is desirable that there should be as many zero correlations as possible in a factor structure. In the ideal case, groups of tests correlate positively with one factor and are uncorrelated with all the others.

d. Simplicity of Pattern.—It is also desirable that there should be as many zero coefficients of group factors as possible in a factor pattern. We shall also speak of the simplicity of each test to stand for the number of vanishing coefficients of group factors in its linear equation. This is complementary to what Thurstone[1] calls the "complexity" of a test, *i.e.*, the number of non-vanishing coefficients of group factors.

It should be especially noted that simplicity of structure does not imply corresponding simplicity of tests, because the latter are expressed as linear functions of factors. The coefficients in such equations become the elements of a factor structure only in the orthogonal case. It is thus impossible to make a factor solution in terms of simple structure and directly infer simplicity of each test.

e. Independence of Factors.—The great statistical simplicity resulting from the use of orthogonal factors makes it highly desirable to employ them whenever possible.

---

[1] *Op. cit.*, p. 155.

f. Statistical Fit.—This is the degree to which the pattern reproduces the original correlations. It is a useful guide in determining the adequacy of the pattern. When the residuals are sufficiently small, we may judge, on statistical grounds, that no further factors are needed. A standard for judging the general adequacy of fit is given by comparing the standard deviation of the residuals with that of a zero correlation for a sample of equal size.

g. Psychological Interpretation.—Some solutions lead to difficult and awkward interpretation of the factors since a large share of the correlations of tests with factors are negative. Since this difficulty may be easily remedied by the rotation of axes it appears to us desirable to do so. The solutions of Kelley and Hotelling, on the other hand, are given special interpretation to conform with the plus and minus entries throughout.

The preceding criteria are not entirely independent of each other, nor completely exhaustive of all possible standards for judging an analysis. We regard them, however, to be of paramount importance as a guide in selecting a factorial solution, out of a possible infinitude, for any given matrix of intercorrelations.

3. SYSTEMS OF FACTORS

In this section we shall consider three orthogonal and two oblique systems of factor axes. These will be listed and briefly described in the following paragraphs:

a. Bi-factor.—This is a system of mutually orthogonal reference axes. The resulting pattern reveals a general factor present in all tests and μ other factors present in varying groups of tests. Hence, this system produces a great simplicity of pattern; in general the simplicity of each test is equal to $(\mu - 1)$, or complexity $= 2$. Thought of in another way, the zero entries in such an orthogonal factor pattern indicate zero correlations of tests with appropriate factors, and reveal a very simple structure for the μ group factors. We shall use u to denote the general factor and $\alpha, \beta, \gamma, \ldots, \mu$ to denote the μ group factors in order to conform to the notation of our Preliminary Reports.[1]

b. Centroid.—This is the primary system of reference axes in multiple factor analysis. The axes are mutually orthogonal passing through "average" positions of the test vectors. Plus and minus entries will appear for all but the first of the M group factors in the pattern or structure. In general there will be $M \cdot n$ numbers different from zero in the pattern or structure and hence this system of axes

[1] *Preliminary Reports on Spearman-Holzinger Unitary Trait Study*, Nos. 1–9.

yields neither simplicity of pattern nor simplicity of structure. The notation for these axes shall be C_1, C_2, C_3, . . . , C_M instead of I, II, III, . . . , M as given by Thurstone.[1]

c. Multiple Orthogonal.—We shall use the name, "multiple orthogonal" to refer to the system of orthogonal reference axes obtained from the centroid system by an orthogonal transformation which yields the following properties: These axes pass as nearly as possible (subject to the condition that they be orthogonal) through clusters of test vectors and hence may show good geometric fit. In the resulting pattern or structure, since this is an orthogonal case, there will thus be more negligible or zero entries than in the centroid pattern. This system then approximates simplicity of structure and pattern. We shall employ the notation Z_1, Z_2, Z_3, . . . , Z_M to denote such axes.

d. Simple.—This system of oblique reference vectors may be obtained by a transformation on the centroid system. Thurstone refers to such a system as, "oblique reference vectors." The simple axes are the perpendiculars to hyperplanes passing through points representing groups of tests and the origin. This is done to obtain simplicity of structure, since all tests in a plane will have zero correlations with the perpendicular to that plane. Oblique simple axes in general do not yield a simple pattern. The notation for these axes will be V_A, V_B, V_C, . . . , V_M instead of Λ_A, Λ_B, Λ_C, . . . , Λ_M as used by Thurstone.

e. Primary.—We shall use this term in the same sense Thurstone does. These axes are generally oblique. They are formed by the intersections of $(M - 1)$ of the M hyperplanes which gave the simple axes. The resulting factor structure will not be simple because the primary and simple axes are generally distinct. The system, however, yields a simplicity of pattern as a consequence of the simplicity of structure with respect to the simple axes. All tests in a hyperplane (which have zero correlations with the simple axis perpendicular to that hyperplane) can be expressed linearly in terms of the $(M - 1)$ primary factors lying in that hyperplane. Hence the simplicity of each test is at least one. Since the primary structure is not simple in character it is difficult to identify the factors thereby. It is possible, however, to name the primary factors from the structure of the simple factors as is done by Thurstone.[2] The notation for these axes will be T_A, T_B, T_C, . . . , T_M.

[1] *Loc. cit.*

[2] *Op. cit.*, Chapter VI.

The notation for the five systems of factors given above is summarized as follows:

Bi-factor..... $u, \alpha, \beta, \gamma, \ldots, \mu$;
Centroid..... $C_1, C_2, C_3, \ldots, C_M$;
Multiple orthogonal..... $Z_1, Z_2, Z_3, \ldots, Z_M$;
Simple..... $V_A, V_B, V_C, \ldots, V_M$;
Primary..... $T_A, T_B, T_C, \ldots, T_M$.

Now we may proceed to an hypothetical example to illustrate our methods. We shall set up an example consisting of twelve tests to measure three factors $V_k (k = A, B, C)$ in the most general multiple factor sense. From the intercorrelations of these tests we shall derive structures and patterns for each of the systems listed in this section. Detailed comparisons will be made between the factors obtained from the bi-factor analysis and three of the multiple factor solutions.

4. HYPOTHETICAL EXAMPLE

We assume simple structure as exhibited in the matrix V of Table I, which can be obtained from a centroid matrix by means of the oblique

TABLE I

Test	Matrix V		
	$r_{z_j V_A}$	$r_{z_j V_B}$	$r_{z_j V_C}$
1	.3	..	..
2	.3	..	..
3	.4	..	..
4	.5	..	..
5	..	.6	..
6	..	.7	..
7	..	.5	..
8	..	.7	..
9	..	..	.2
10	..	..	.3
11	..	..	.4
12	..	..	.5

transformation of Table II. We start with dependent factors as the most general solution of multiple factor analysis since, "the structure of the factorial matrix can be simplified by introducing oblique reference vectors."[1] The matrix V was obtained from an assumed centroid matrix F by means of the transformation G, hence

[1] Thurstone: *Op. cit.*, p. 154.

$$V = FG. \tag{5}$$

By multiplying both sides of this equation by G^{-1} on the right we obtain

$$F = VG^{-1}. \tag{6}$$

TABLE II

	Matrix G		
	V_A	V_B	V_C
C_1	.3	.4	.3
C_2	−.3	.9	−.5
C_3	.9	−.1	−.8

The inverse of G exists because G is non-singular, its determinant being $|G| = -.741$.

Computing the inverse of G, we have

$$G^{-1} = \begin{vmatrix} 1.04 & -.39 & .63 \\ .93 & .69 & -.08 \\ 1.05 & -.53 & -.53 \end{vmatrix}. \tag{7}$$

We perform the indicated matrix multiplication in equation (6) and

TABLE III

Test	Matrix F			$\sqrt{\sum_{i=1}^{3} r^2_{z_j C_i}}$	Direction cosines		
	C_1	C_2	C_3		C_1	C_2	C_3
1	.31	−.12	.19	.383	.81	−.31	.50
2	.31	−.12	.19	.383	.81	−.31	.50
3	.42	−.16	.25	.514	.82	−.31	.49
4	.52	−.20	.32	.642	.81	−.31	.50
5	.56	.41	−.05	.696	.80	.59	−.07
6	.65	.48	−.06	.810	.80	.59	−.07
7	.47	.35	−.04	.587	.80	.60	−.07
8	.65	.48	−.06	.810	.80	.59	−.07
9	.21	−.11	−.11	.261	.80	−.42	−.42
10	.32	−.16	−.16	.392	.82	−.41	−.41
11	.42	−.21	−.21	.514	.82	−.41	−.41
12	.53	−.27	−.27	.653	.81	−.41	−.41

obtain the centroid pattern (matrix F), which is printed in the first three columns of Table III.

From the matrix F we can obtain, by means of equation (4), intercorrelations between the twelve tests. These correlations are presented in Table IV. When the table of intercorrelations is factored by the centroid method,[1] the results check with the values of Table III within errors of rounding. Hence, for the sake of economy of space and unnecessary additional tables, we shall consider matrix F of Table III as obtained from the matrix of intercorrelations.

TABLE IV.—INTERCORRELATIONS OF TWELVE VARIABLES

	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	.15											
3	.20	.20										
4	.25	.25	.33									
5	.11	.11	.16	.19								
6	.13	.13	.18	.22	.56							
7	.10	.10	.13	.16	.41	.48						
8	.13	.13	.18	.22	.56	.66	.48					
9	.06	.06	.08	.10	.08	.09	.06	.09				
10	.09	.09	.12	.15	.12	.14	.10	.14	.10			
11	.12	.12	.16	.19	.16	.18	.13	.18	.13	.20		
12	.15	.15	.20	.24	.20	.23	.17	.23	.17	.26	.34	

Each row of matrix F represents the three coördinates of the respective test (considered as a point in a 3-space) referred to the centroid frame of reference—C_1, C_2, C_3. If we consider each test as a vector (the directed line from the origin to the point represented by the three coördinates) then each row of F represents numbers proportional to the direction cosines of the respective test vectors. We can normalize these test vectors, that is, make their lengths equal to one, by dividing each row of F by the square root of the sum of the squares of the entries in the respective rows. Thus normalized, the entries in each row are the actual direction cosines of the test vectors. The direction cosines are given in the last three columns of Table III.

Inasmuch as the test vectors are normalized, their termini can be represented as points on the surface of a unit sphere. In Figure 1 we have plotted the points representing the twelve test vectors. Since the tests cluster around three distinct groups (as is evident from the direction cosines of the tests), we have represented the twelve tests

[1] Thurstone: *Op. cit.*, Chapter III.

simply by three points, A, B and C. Point A is the centroid of tests 1, 2, 3, 4; point B is the centroid of tests 5, 6, 7, 8; and point C is the centroid of tests 9, 10, 11, 12. The coördinates of the three points are as follows:

$$\begin{aligned} A&: \ (.81, -.31, \ .50), \\ B&: \ (.80, \ .59, -.07), \\ C&: \ (.81, -.41, -.41). \end{aligned} \tag{8}$$

The frame of reference is set up so that the plane of the paper coincides

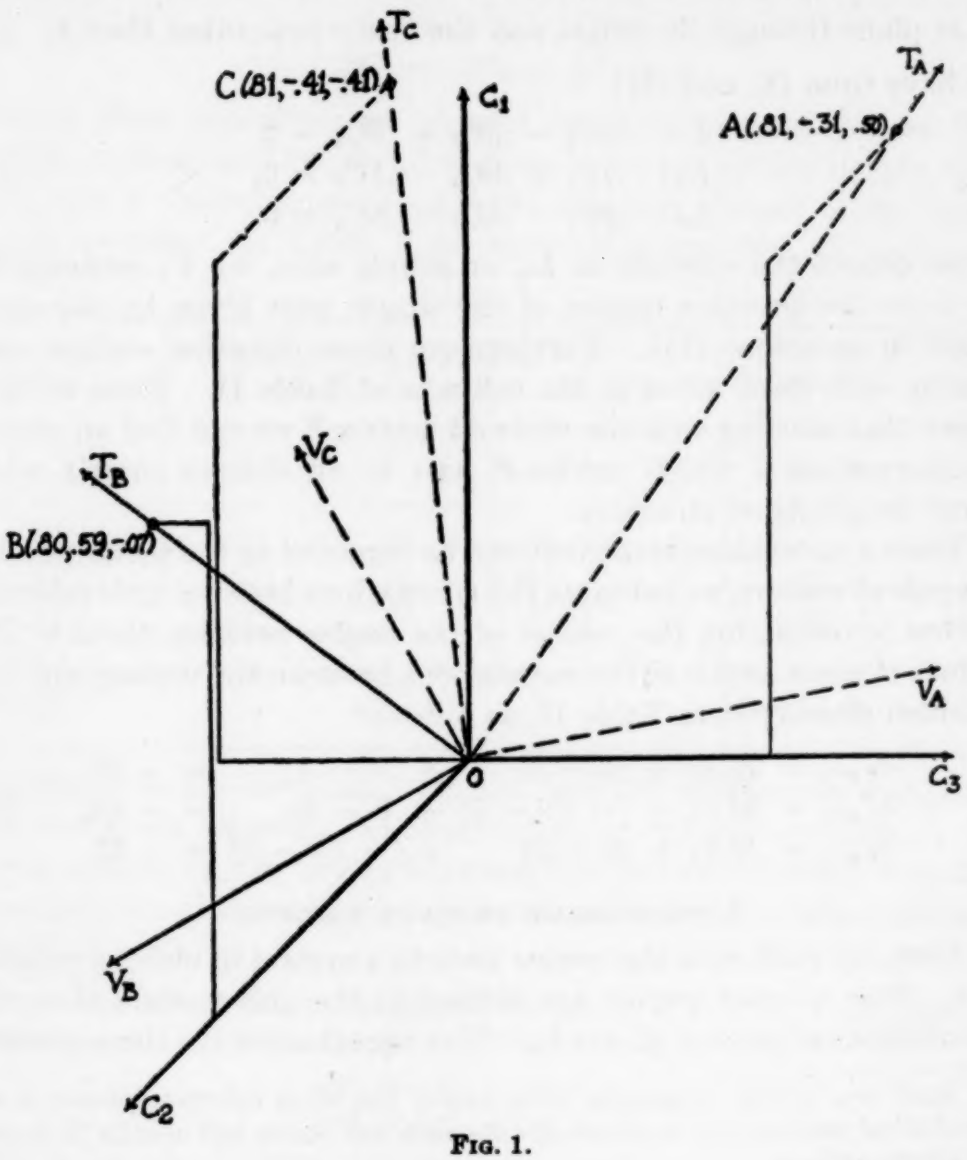


FIG. 1.

with the plane formed by the centroid axes C_1 and C_3. The axis C_2 is then perpendicular to the plane of the paper. The diagram is

shown with the surface of the sphere removed so that the origin and the positive ends of the three axes are visible.

Now we may pass three reference planes through the origin O and pairs of the points A, B, C so that each of the test vectors lies in two of the three planes.[1] The general equation of a plane through the origin is

$$\lambda C_1 + \mu C_2 + \nu C_3 = 0, \tag{9}$$

where λ, μ, ν are the direction cosines of the normal to the plane. Indicating the three planes by L_k, $k = A, B, C$, where

$L_k \equiv$ plane through the origin and the two points other than k, (10)

we have from (8) and (9):

$$\begin{aligned} L_A&: \ .3C_1 - .3C_2 + .9C_3 = 0, \\ L_B&: \ .4C_1 + .9C_2 - .1C_3 = 0, \\ L_C&: \ .3C_1 - .5C_2 - .8C_3 = 0. \end{aligned} \tag{11}$$

If we denote the normals to L_k, or simple axes, by V_k respectively, we have the direction cosines of the simple axes given by the coefficients in equations (11). Furthermore these direction cosines agree exactly with those given in the columns of Table II. Thus we have shown that starting with the centroid matrix F we can find an oblique transformation G which carries F into V, an oblique matrix which shows simplicity of structure.

Since a correlation coefficient can be regarded as the scalar product of a pair of vectors, we compute the correlations between unit reference vectors by obtaining the cosines of the angles between them.[2] The cosines of these angles or the correlations between the vectors are then obtained directly from Table II, as follows:

$$\begin{aligned} r_{V_AV_B} &= .3(.4) + .9(-.3) + .9(-.1) = -.24, \\ r_{V_AV_C} &= .3(.3) + (-.3)(-.5) + .9(-.8) = -.48, \\ r_{V_BV_C} &= .4(.3) + .9(-.5) + (-.1)(-.8) = -.25. \end{aligned} \tag{12}$$

5. DEPENDENT PRIMARY FACTORS

Next, we shall refer the twelve tests to a system of oblique primary axes. The primary vectors are defined as the unit vectors along the intersections of pairs of planes L_k. The equations of the three primary

[1] Each test vector happens to lie in two of the three reference planes in our hypothetical problem, but with actual data each test vector will usually lie in only one reference plane.

[2] The cosine of an angle θ between two directed lines whose direction cosines are λ_1, μ_1, ν_1 and λ_2, μ_2, ν_2 is given by

$$\cos\theta = \lambda_1\lambda_2 + \mu_1\mu_2 + \nu_1\nu_2.$$

axes are then given by the simultaneous solution of pairs of equations (11). Denoting these vectors by T_k, where,

$$T_k \equiv \text{line determined by two planes } L_i, L_j \qquad \begin{pmatrix} i,j,k = A,B,C \\ i \neq j \neq k \end{pmatrix}, \tag{13}$$

we have,

$$\begin{aligned} T_A\text{:}\quad & .3C_1 - .5C_2 - .8C_3 = 0, \\ & .4C_1 + .9C_2 - .1C_3 = 0; \\ T_B\text{:}\quad & .3C_1 - .3C_2 + .9C_3 = 0, \\ & .3C_1 - .5C_2 - .8C_3 = 0; \\ T_C\text{:}\quad & .4C_1 + .9C_2 - .1C_3 = 0, \\ & .3C_1 - .3C_2 + .9C_3 = 0. \end{aligned} \tag{14}$$

From these equations we obtain[1] numbers proportional to the direction cosines, and by dividing these numbers by the square root of the sum of their squares we get the direction cosines of the three primary vectors. In Table V we have presented first the numbers

TABLE V

	Numbers proportional to direction cosines			Matrix T		
	τ_A	τ_B	τ_C	T_A	T_B	T_C
C_1	.77	.69	.78	.81	.80	.82
C_2	−.29	.51	−.39	−.31	.59	−.41
C_3	.47	−.06	−.39	.50	−.07	−.41
$\sqrt{\sum_{i=1}^{3} \tau^2_{kC_i}}$	.948	.860	.955			

proportional to the direction cosines of the primary vectors and then, in the last three columns, the direction cosines of these vectors.

The cosines of the angles between the primary vectors, *i.e.*, the intercorrelations between primary vectors, are obtained from matrix T of Table V as follows:

[1] From the theory of analytic geometry of space, we know that the direction cosines λ, μ, ν of a line of intersection of the two planes

$$\begin{aligned} a_1C_1 + b_1C_2 + c_1C_3 &= 0, \\ a_2C_1 + b_2C_2 + c_2C_3 &= 0, \end{aligned}$$

are given by the following ratios:

$$\lambda : \mu : \nu = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} : \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} : \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

$$
\begin{aligned}
r_{T_AT_B} &= .81(.80) + .59(-.31) + .50(-.07) = .43, \\
r_{T_AT_C} &= .81(.82) + (-.31)(-.41) + .50(-.41) = .59, \\
r_{T_BT_C} &= .80(.82) + .59(-.41) + (-.07)(-.41) = .44.
\end{aligned} \tag{15}
$$

We refer the tests to the primary axes by multiplying matrix F of Table III by matrix T of Table V. The resulting T structure is presented in Table VI. As pointed out in Section 3, this table illus-

TABLE VI

Test	T structure		
	$r_{z_jT_A}$	$r_{z_jT_B}$	$r_{z_jT_C}$
1	.38	.16	.23
2	.38	.16	.23
3	.51	.22	.31
4	.64	.28	.38
5	.30	.69	.31
6	.35	.81	.36
7	.25	.59	.26
8	.35	.81	.36
9	.15	.11	.26
10	.23	.17	.39
11	.30	.23	.52
12	.38	.28	.66

trates that the primary structure does not furnish so useful a basis for naming the factors as the structure of the simple factors. The positive values in Table VI do not reveal the identifying groups of tests so clearly as do the values of Table I.

6. MULTIPLE ORTHOGONAL FACTORS

Another multiple factor solution of interest is one with independent factors. We impose the condition of orthogonality and attempt also to get simplicity of structure. We rotate the centroid axes two at a time to obtain our new reference axes.

In Figure 2, we have plotted the twelve points, whose centroid coördinates are given in matrix F, in the C_1, C_2-plane. The most effective transformation is obtained by passing a new set of orthogonal axes Y_1, Y_2 as near to clusters of tests as possible. In this case the new axes make an angle of $-45°$ with the old axes. The equations of transformation from one set of rectangular axes to another in the plane are given by

$$
\begin{aligned}
C_1 &= Y_1 \cos\theta - Y_2 \sin\theta, \\
C_2 &= Y_1 \sin\theta + Y_2 \cos\theta,
\end{aligned} \tag{16}
$$

where θ is the angle from the positive old axis to the corresponding positive new axis. Applying transformation (16) with $\theta = -45°$

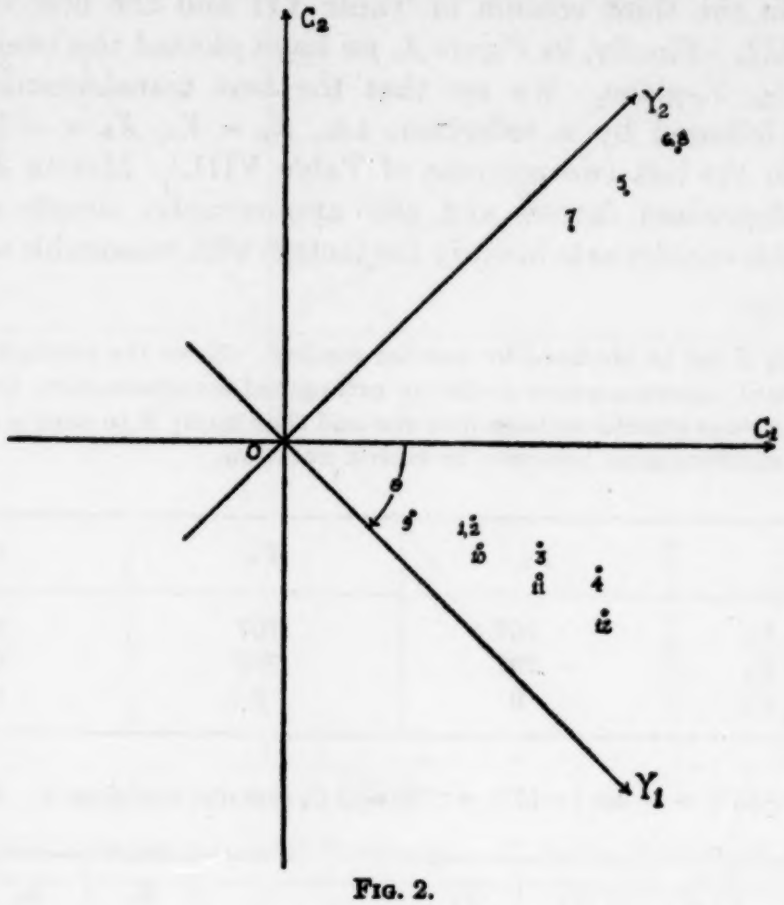


FIG. 2.

to the first two columns of matrix F we obtain the first two columns of Table VII.

TABLE VII.—PARTIAL TRANSFORMATION OF MATRIX F

Test	Y_1	Y_2	Y_3
1	.304	.134	−.019
2	.304	.134	−.019
3	.410	.184	−.031
4	.509	.226	−.030
5	.106	.686	−.102
6	.120	.799	−.118
7	.085	.580	−.082
8	.120	.799	−.118
9	.226	.071	−.220
10	.339	.113	−.326
11	.445	.148	−.427
12	.566	.184	−.546

Next, we plot C_3 against Y_1 in Figure 3. Making a transformation to new orthogonal axes Z_1, Y_3 at an angle $\theta = 35°$ with the old axes, we obtain the third column of Table VII and the first column of Table VIII. Finally, in Figure 4, we have plotted the twelve points in the Y_2, Y_3-plane. We see that the best transformation is the identity followed by a reflection, *i.e.*, $Z_2 = Y_2$, $Z_3 = -Y_3$. Thus we obtain the last two columns of Table VIII.[1] Matrix Z exhibits three independent factors and also approximates simple structure. Such a table enables us to identify the factors with reasonable assurance.

---

[1] Matrix Z can be obtained by another method. Since the product of a series of orthogonal transformations is also an orthogonal transformation, we can combine the previous transformations into one and then apply it to matrix F to get Z. The first transformation becomes, in matrix notation,

		Y_1	Y_3	C_3
	C_1	.707	.707	0
S_1:	C_2	−.707	.707	0
	C_3	0	0	1

since cos (−45°) = − sin (−45°) = .707 and C_3 remains unchanged. Similarly,

		Z_1	Y_2	Y_3
	Y_1	.819	0	−.574
S_2:	Y_2	0	1	0
	C_3	.574	0	.819

,

		Z_1	Z_2	Z_3
	Z_1	1	0	0
S_3:	Y_2	0	1	0
	Y_3	0	0	−1

,

since cos 35° = .819, sin 35° = .574, and cos 0° = 1, sin 0° = 0. The product of the three transformations is given by

		Z_1	Z_2	Z_3
	C_1	.579	.707	.406
$(S_1S_2S_3)$:	C_2	−.579	.707	−.406
	C_3	.574	0	−.819

.

Upon multiplying F by $(S_1S_2S_3)$ we have Z, *i.e.*, $Z = F(S_1S_2S_3)$.

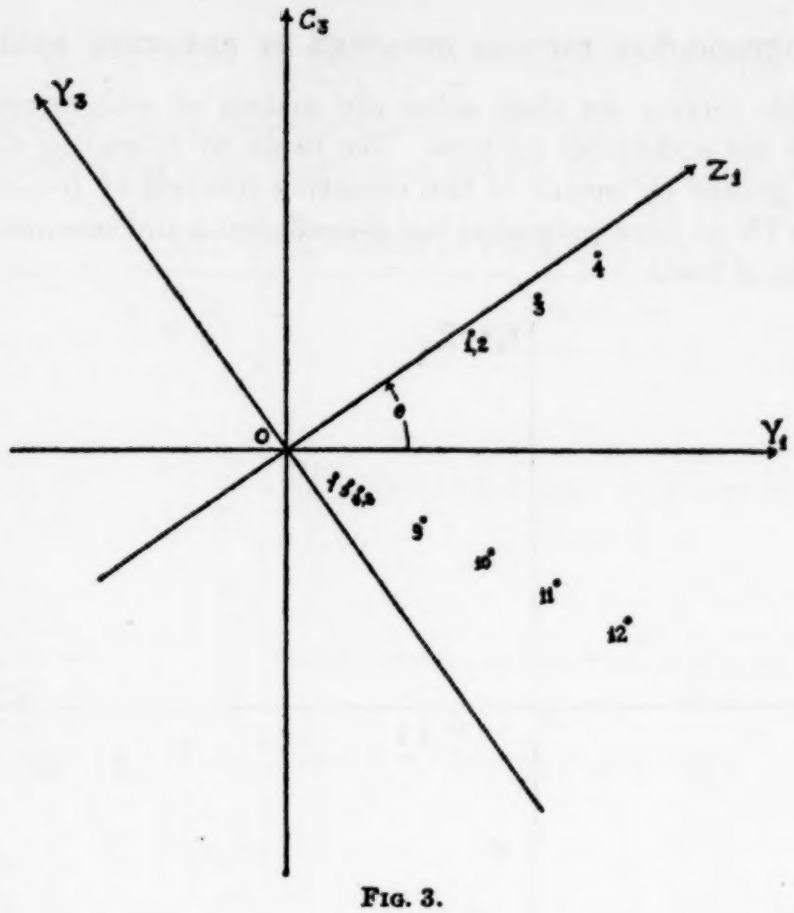


FIG. 3.

TABLE VIII.—COEFFICIENTS OF ORTHOGONAL MULTIPLE FACTOR PATTERN

Test	Matrix Z		
	Z_1	Z_2	Z_3
1	.358	.134	.019
2	.358	.134	.019
3	.479	.184	.031
4	.601	.226	.030
5	.058	.686	.102
6	.064	.799	.118
7	.047	.580	.082
8	.064	.799	.118
9	.122	.071	.220
10	.186	.113	.326
11	.244	.148	.427
12	.309	.184	.546

7. ORTHOGONAL FACTORS OBTAINED BY BI-FACTOR ANALYSIS

In this section we shall solve the matrix of intercorrelations of Table IV for a bi-factor pattern. We begin by allocating the twelve tests to groups by means of the objective method of *B*-coefficients.[1] In Table IX we have presented the *B*-coefficients and comments on the allocation of tests.

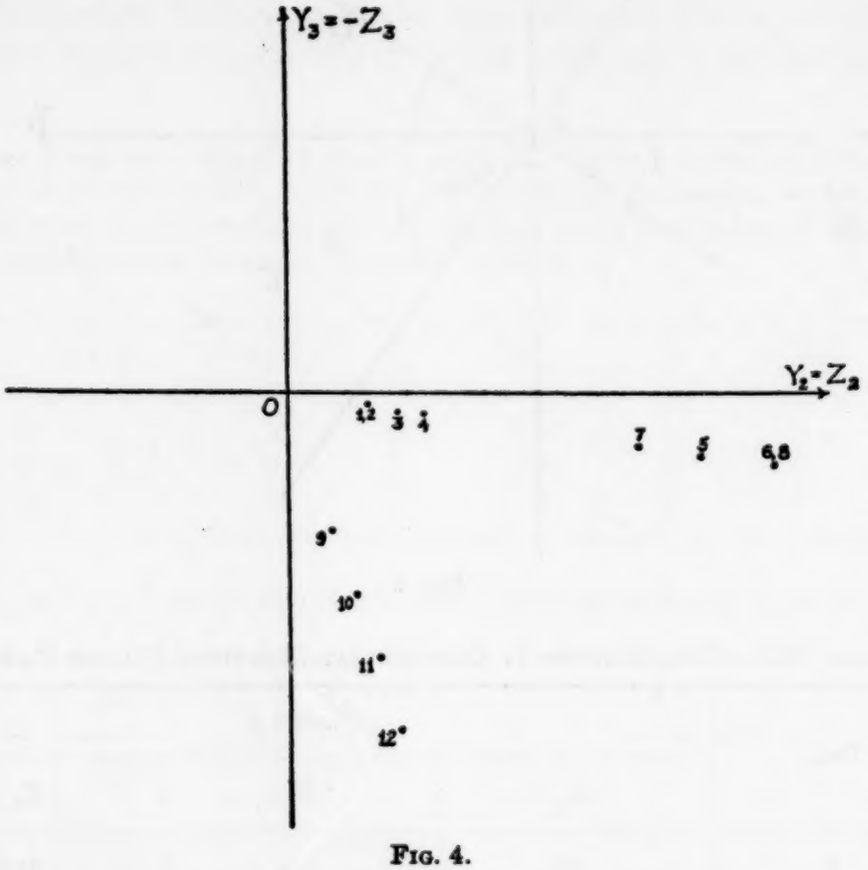


FIG. 4.

We now obtain[2] the correlations of the tests with the general factor u, which appear as the entries in the first column of Table X. Next, we get the residual correlations with u removed as presented below the principal diagonal in Table XI. Then we obtain[2] the correlations of tests with the group factors α, β, and γ, which appear in Table X. The final residuals with the group factors out are given above the diagonal opposite the respective residuals with only u out in Table XI. A solution of the type exhibited in Table X enables one to identify the factors without any difficulty.

---

[1] *Preliminary Report on Spearman-Holzinger Unitary Trait Study*, No. 7.
[2] *Ibid.*

8. PATTERNS IN TERMS OF DEPENDENT FACTORS

Since we propose to make a comparison of factors from the linear descriptions of the tests, it will be necessary to obtain factor patterns

TABLE IX.—ALLOCATION OF TESTS INTO GROUPS

$B(x_1, x_2, \ldots, x_k)$	B	Remarks
$B(6, 8)$.....	2.82	
$B(6, 8, 5)$.....	3.14	
$B(6, 8, 5, 7)$.....	3.60	(β)
$B(6, 8, 5, 7, 12)$.....	2.60	Reject tests 12 and 4 because of large drop.
$B(6, 8, 5, 7, 4)$.....	2.56	
$B(11, 12)$.....	1.90	
$B(11, 12, 10)$.....	1.73	
$B(11, 12, 10, 4)$.....	1.39	Reject tests 4 and 3 because of large drop.
$B(11, 12, 10, 3)$.....	1.32	
$B(11, 12, 10, 9)$.....	1.46	(γ)
$B(11, 12, 10, 9, 4)$...	1.24	Reject test 4 because of large drop.
$B(3, 4)$.....	1.84	
$B(3, 4, 2)$.....	1.69	
$B(3, 4, 2, 1)$.....	1.65	(α)

TABLE X.—COEFFICIENTS OF BI-FACTOR PATTERN

Test	u	α	β	γ
1	.29	.26	...	...
2	.29	.26	...	...
3	.40	.32	...	...
4	.48	.43	...	...
5	.39	...	.56	...
6	.45	...	.68	...
7	.33	...	.49	...
8	.45	...	.68	...
9	.20	...	...	.16
10	.31	...	...	.24
11	.40	...	...	.33
12	.51	...	...	.43

for the oblique solutions. We shall derive a V pattern and a T pattern in this section.

TABLE XI.—FINAL RESIDUAL CORRELATIONS

	1	2	3	4	5	6	7	8	9	10	11	12
1	...	.00	.00	.00								
2	.07	...	.00	.00								
3	.08	.08	...	.00								
4	.11	.11	.14									
5	.00	.00	.00	.00	...	.00	.01	.00				
6	.00	.00	.00	.00	.38	...	.00	.00				
7	.00	.00	.00	.00	.28	.33	.....	.00				
8	.00	.00	.00	.00	.38	.46	.33					
9	.00	.00	.00	.00	.00	.00	−.01	.00	...	.00	.00	.00
10	.00	.00	.00	.00	.00	.00	.00	.00	.04	...	.00	.00
11	.00	.00	.00	.00	.00	.00	.00	.00	.05	.08	...	.00
12	.00	.00	.00	.00	.00	.00	.00	.00	.07	.10	.14	

In the case of the V pattern, the tests x_j are expressed linearly in terms of the factors V_k,

$$x_j = v_{jA}V_A + v_{jB}V_B + v_{jC}V_C, \qquad (j = 1, 2, 3, \ldots, 12), \quad (17)$$

where the thirty-six coefficients v_{jk} are to be determined. We can obtain twelve sets of three equations each for the correlations of the tests with the three V-factors as follows:

$$\begin{aligned} r_{x_jV_A} &= v_{jA} + v_{jB}r_{V_AV_B} + v_{jC}r_{V_AV_C}, \\ r_{x_jV_B} &= v_{jA}r_{V_BV_A} + v_{jB} + v_{jC}r_{V_BV_C}, \\ r_{x_jV_C} &= v_{jA}r_{V_CV_A} + v_{jB}r_{V_CV_B} + v_{jC}. \end{aligned} \qquad (18)$$

The values of $r_{x_jV_k}$ are given in Table I, and the intercorrelations of factors $r_{V_kV_l}(k, l = A, B, C)$ are given in equations (12). Hence for every j, the system (18) can be considered as three simultaneous equations in the three unknowns v_{jA}, v_{jB}, and v_{jC}. Thus we can solve for the v_{jk} and substitute their values in equations (17) to get the twelve linear expressions of the tests in terms of the dependent factors V_A, V_B, and V_C. We have presented the coefficients of this pattern in Table XII.

In a similar manner, we can obtain a T pattern from the T structure and intercorrelations of the T-factors. The linear expressions for the tests in terms of these factors,

$$x_j = t_{jA}T_A + t_{jB}T_B + t_{jC}T_C, \qquad (j = 1, 2, 3, \cdots, 12), \quad (19)$$

involve the coefficients t_{jk} which are to be determined. The twelve

TABLE XII.—COEFFICIENTS OF V-FACTOR PATTERN

Test	V_A	V_B	V_C
1	.475	.182	.274
2	.475	.182	.274
3	.634	.243	.365
4	.792	.304	.456
5	.365	.780	.370
6	.426	.910	.432
7	.304	.650	.308
8	.426	.910	.432
9	.182	.123	.318
10	.274	.185	.478
11	.365	.247	.637
12	.456	.308	.796

sets of three equations each for the correlations of the tests with the three T-factors are given by the equations,

$$\begin{aligned} r_{x_i T_A} &= t_{jA} + t_{jB} r_{T_A T_B} + t_{jC} r_{T_A T_C}, \\ r_{x_i T_B} &= t_{jA} r_{T_B T_A} + t_{jB} + t_{jC} r_{T_B T_C}, \\ r_{x_i T_C} &= t_{jA} r_{T_C T_A} + t_{jB} r_{T_C T_B} + t_{jC}. \end{aligned} \tag{20}$$

Knowing the values of $r_{x_i T_A}$, $r_{x_i T_B}$, $r_{x_i T_C}$ from Table VI, and the values of $r_{T_A T_B}$, $r_{T_A T_C}$, $r_{T_B T_C}$ from equations (15), we can solve for the t_{jA}, t_{jB}, t_{jC} for each j from the corresponding system (20) of three equations. When we have found the values of t_{jk}, we substitute them into equations (19) to get the factor pattern.

In Table XIII we have presented the coefficients of the T pattern. As pointed out in Paragraph (*e*) of Section 3, Table XIII exhibits a simplicity of pattern. Except for rounding errors in computation, we should expect eight zeros in each column of this table.

9. COMPARISON OF FACTORIAL SOLUTIONS

In this section we shall show the relationships between the factors obtained by the bi-factor analysis and several multiple factor analyses. The general procedure shall be to equate two linear expressions for the same test from different factor patterns and from a sufficient number of such equalities to solve for one group of factors in terms of the other.

If we have n linearly independent non-homogeneous equations in n unknowns we are able to solve for these unknowns. We shall

TABLE XIII.—COEFFICIENTS OF T-FACTOR PATTERN

Test	T_A	T_B	T_C
1	.376	−.007	.011
2	.376	−.007	.011
3	.502	−.002	.015
4	.636	.005	.002
5	.000	.686	.008
6	.000	.808	.005
7	−.006	.591	.004
8	.000	.808	.005
9	−.004	−.005	.264
10	.000	−.002	.391
11	−.011	.004	.525
12	−.012	−.010	.671

illustrate the algebraic procedure of comparing factors by means of our hypothetical example of twelve tests. In this case we shall have three equations in three unknowns.

The following groups of four tests each, (1, 2, 3, 4), (5, 6, 7, 8), and (9, 10, 11, 12), involve fewer than four factors in the linear expressions of the tests, and hence the four forms in each group are linearly dependent.[1] We shall refer to the centroids of these test groups by x_A, x_B, and x_C respectively. Then for each of the multiple factor patterns we can write three equations in three factors:

$$\begin{aligned} x_A &= f_{AA}F_A + f_{AB}F_B + f_{AC}F_C, \\ x_B &= f_{BA}F_A + f_{BB}F_B + f_{BC}F_C, \\ x_C &= f_{CA}F_A + f_{CB}F_B + f_{CC}F_C, \end{aligned} \tag{21}$$

where the $f_{kl}(k, l = A, B, C)$ are constants, and F_k are the respective factors in the different patterns. The functions (21) are linearly independent since the determinant[2] of coefficients in each case is different from zero.[3]

The bi-factor pattern yields the following three linearly independent equations,[4]

[1] For an elementary, direct proof see Dickson: *Op. cit.*, Corollary 1 to Theorem 13, p. 59.

[2] *Infra.*, determinants in systems of equations (23), (24), and (25).

[3] Dickson: *Op. cit.*, Theorem 9, p. 55.

[4] *Ibid.*, Corollary 2 to Theorem 13, p. 60.

$$\begin{aligned} x_A &= .365u + .318\alpha, \\ x_B &= .405u + .603\beta, \\ x_C &= .355u + .290\gamma. \end{aligned} \tag{22}$$

In the case of the V-factor pattern, equations (21) reduce to

$$\begin{aligned} x_A &= .594V_A + .228V_B + .342V_C, \\ x_B &= .380V_A + .813V_B + .386V_C, \\ x_C &= .319V_A + .216V_B + .557V_C. \end{aligned} \tag{23}$$

Similarly, for the T-factor pattern, equations (21) become

$$\begin{aligned} x_A &= .473T_A - .003T_B + .010T_C, \\ x_B &= -.002T_A + .723T_B + .006T_C, \\ x_C &= -.007T_A - .003T_B + .463T_C. \end{aligned} \tag{24}$$

Finally, for the Z-factor pattern, the three linearly independent equations follow,

$$\begin{aligned} x_A &= .449Z_1 + .170Z_2 + .025Z_3, \\ x_B &= .058Z_1 + .716Z_2 + .105Z_3, \\ x_C &= .215Z_1 + .129Z_2 + .380Z_3. \end{aligned} \tag{25}$$

By equating the respective equations of sets (22) and (23) we have the three following equations,

$$\begin{aligned} .594V_A + .228V_B + .342V_C &= .365u + .318\alpha, \\ .380V_A + .813V_B + .386V_C &= .405u + .603\beta, \\ .319V_A + .216V_B + .557V_C &= .355u + .290\gamma. \end{aligned} \tag{26}$$

We may consider the set (26) as representing three linearly independent equations in the three unknowns V_A, V_B, V_C, with the right-hand members as known quantities. Then we have sufficient conditions for the solution[1] of V_A, V_B, V_C in terms of u, α, β, γ. Ordinary methods[2] for solving these equations give us,

[1] *Ibid.*, Theorem 16, p. 63.

[2] Solving by means of determinants, we have,

$$V_A = \frac{\begin{vmatrix} (.365u + .318\alpha) & .228 & .342 \\ (.405u + .603\beta) & .813 & .386 \\ (.355u + .290\gamma) & .216 & .557 \end{vmatrix}}{\begin{vmatrix} .594 & .228 & .342 \\ .380 & .813 & .386 \\ .319 & .216 & .557 \end{vmatrix}}$$

$$= \frac{.36947(.365u + .318\alpha) - .05312(.405u + .603\beta) - .19004(.355u + .290\gamma)}{.13865}$$

$$= .3309u + .8474\alpha - .2310\beta - .3975\gamma, \text{ etc.}$$

$$\begin{aligned} V_A &= .3309u + .8474\alpha - .2310\beta - .3975\gamma, \\ V_B &= .1604u - .2030\alpha + .9645\beta - .2077\gamma, \\ V_C &= .3856u - .4066\alpha - .2417\beta + .8289\gamma. \end{aligned} \tag{27}$$

These equations are the exact algebraic relationships between the V-factors obtained by multiple factor analysis and the factors obtained by bi-factor analysis. Equations (27) indicate that the statistically dependent factors V_A, V_B, and V_C are most like the independent factors α, β, and γ respectively. The correspondence, however, is not very close since there is some overlapping of all the orthogonal factors in each of the oblique factors. The correlations $r_{x_k x_l}$ $(k, l = A, B, C)$ and the communalities of x_k for the two solutions should be equal in order to obtain such relationships in standard form.

Let us denote the coefficients in equations (27) as follows:

$$\begin{aligned} v_{Au} &= .3309, & v_{A\alpha} &= .8474, & v_{A\beta} &= -.2310, & v_{A\gamma} &= -.3975; \\ v_{Bu} &= .1604, & v_{B\alpha} &= -.2030, & v_{B\beta} &= .9645, & v_{B\gamma} &= -.2077; \\ v_{Cu} &= .3856, & v_{C\alpha} &= -.4066, & v_{C\beta} &= -.2417, & v_{C\gamma} &= .8289. \end{aligned} \tag{28}$$

Then, if we multiply any one of the factors V_k by x_j, sum over the number of cases N, and divide by N, we have,

$$r_{x_j V_k} = v_{ku} r_{x_j u} + v_{k\alpha} r_{x_j \alpha} + v_{k\beta} r_{x_j \beta} + v_{k\gamma} r_{x_j \gamma}, \qquad \begin{pmatrix} k = A, B, C; \\ j = 1, 2, \cdots, 12 \end{pmatrix}. \tag{29}$$

This is a symbolic equation representing thirty-six such equations; for each value of $k = A, B, C$ we have twelve equations.

TABLE XIV.—V-FACTOR STRUCTURE DERIVED FROM BI-FACTOR STRUCTURE

Test	$r_{x_j V_A}$	$r_{x_j V_B}$	$r_{x_j V_C}$
1	.316	−.006	.006
2	.316	−.006	.006
3	.404	−.001	.024
4	.523	−.010	.010
5	.000	.603	.015
6	−.008	.728	.009
7	−.004	.526	.009
8	−.008	.728	.009
9	.003	−.001	.210
10	.007	.000	.318
11	.001	−.004	.428
12	−.002	−.008	.553

We may thus write the V-factor structure as obtained in equations (29) from the bi-factor structure of Table X. This derived factor structure is given in Table XIV. On comparing this table with the original V-factor structure (Table I) we see that there are only slight discrepancies due to rounding in computation.

In a similar manner, we may equate respective members of sets (22) and (24) and thus solve for the T-factors in terms of u, α, β, and γ. The result is

$$\begin{aligned} T_A &= .7587u + .6721\alpha + .0052\beta - .0133\gamma, \\ T_B &= .5558u + .0018\alpha + .8340\beta - .0052\gamma, \\ T_C &= .7821u + .0102\alpha + .0055\beta + .6261\gamma. \end{aligned} \tag{30}$$

We also have

$$r_{z_j T_k} = t_{ku} r_{z_j u} + t_{k\alpha} r_{z_j \alpha} + t_{k\beta} r_{z_j \beta} + t_{k\gamma} r_{z_j \gamma}, \qquad \begin{pmatrix} k = A, B, C; \\ j = 1, 2, \cdots, 12 \end{pmatrix}, \tag{31}$$

where the t_{ku}, $t_{k\alpha}$, $t_{k\beta}$, and $t_{k\gamma}$ are the constants of equations (30). Equations (31) can be used to reproduce the T-factor structure and hence as a check on the algebraic work of obtaining relations (30).

Finally, we may compare the orthogonal multiple factor pattern with the bi-factor pattern by equating respective members of systems (22) and (25). The relationships follow,

$$\begin{aligned} Z_1 &= .6175u + .7307\alpha - .3291\beta - .0001\gamma, \\ Z_2 &= .4521u + .0017\alpha + .8855\beta - .1178\gamma, \\ Z_3 &= .4309u - .4961\alpha - .1147\beta + .8032\gamma. \end{aligned} \tag{32}$$

Before we analyze these algebraic relationships we shall show how to check the factor structure (or pattern in this case) obtainable from equations (32). We have

$$r_{z_j z_k} = z_{ku} r_{z_j u} + z_{k\alpha} r_{z_j \alpha} + z_{k\beta} r_{z_j \beta} + z_{k\gamma} r_{z_j \gamma}, \qquad \begin{pmatrix} k = A, B, C; \\ j = 1, 2, \cdots, 12 \end{pmatrix}, \tag{33}$$

where the z_{ku}, $z_{k\alpha}$, $z_{k\beta}$, and $z_{k\gamma}$ are the constants in equations (32). By means of equations (33) and the bi-factor structure we get a derived Z-factor structure.

From equations (32) we can compare the orthogonal factors of a multiple factor analysis with those of a bi-factor analysis. Factor Z_1 is most like α with a good deal of the general factor u. Factor Z_2 is almost identical with the bi-factor β but includes some of u. The multiple factor Z_3, although not so significant as the preceding factors, is a great deal like γ with some of u and somewhat different from α.

DEXTRALITY QUOTIENTS OF SEVEN-YEAR-OLDS IN TERMS OF HAND USAGE*

WENDELL JOHNSON AND DOROTHY M. DAVIS

State University of Iowa

I. INTRODUCTION

This study involves an attempt to improve the Iowa Hand Usage Test as constructed by Johnson and Duke;[1] to establish, by means of the improved test, hand usage dextrality quotient norms for seven-year-olds; and to ascertain the effect of age upon the hand usage dextrality quotient by comparing the seven-year-old norms with the tentative six-year-old norms established by Johnson and Duke.

Johnson and Duke originally constructed a test of twenty-seven items which preliminary observations had demonstrated to be indicative of hand preference. This test was scored in terms of the dextrality quotient (DQ) for each child. "The DQ is to be defined as the percentage of the total achievement involved in any test of handedness which is to be credited to the right hand."[1] The formula for obtaining the DQ is $\frac{R + \frac{B}{2}}{N}$ in which R represents the number of activities in which the right hand predominates, and B the activities in which both hands take part, neither predominating. N is the number of activities performed.

Johnson and Duke tentatively standardized the test for six-year-olds by testing fifty six-year-old children under controlled conditions.

It was thought worth while, before establishing additional age norms, to discover the effect of increasing the number of items to forty and of recording only the first two responses—and thus the same number of responses—for each activity.

Also, it was proposed to improve the test by removing any items failing to differentiate between the upper and lower quartiles of the distribution for the whole test, such differentiation being regarded as a criterion of validity.

The forty items selected for the revised form of the test included twenty-six of the twenty-seven original ones, shuffling cards being

* From the Iowa Child Welfare Research Station and the State University of Iowa Speech Clinic. All data and materials pertinent to this study are on file in the State University of Iowa Library with the manuscript copy of this report.

deleted because it had been found an invalid item in the original research. Pointing to cards was separated from pointing to blocks, and these were then treated as separate items. The additional items chosen were believed to fit the testing situation so that it could be organized and conducted as a unit.

II. THE IOWA HAND USAGE TEST (REVISED FORM)

The materials needed for the test are: A room with a window having a shade which can be easily manipulated by a seven-year-old child; a desk of the type from which articles are taken with equal ease from both sides, *i.e.*, the open shelf type, and in which the articles to be removed are not observed; a pencil; a small pencil sharpener of the type held in one hand; a tablet of paper of customary school size; a pencil with eraser attached; a box of crayons; scissors in which the finger holes are the same on both sides; five regular flash cards; a blackboard; a piece of chalk; eight one-and-one-quarter inch cubes; a fountain pen with regulation screw top and side lever for filling; a bottle of ink; and a large picture book with pages which can be readily handled.

Following is the procedure for giving the test, as it was finally revised in this study. The starred items were not included in the Johnson-Duke form of the test; they were added in order to increase the number of items to forty.

Before calling the child, the tablet, pencil, pencil sharpener, and crayons are placed in the desk. As soon as the child enters the room the following directions are given:

"Will you please **pull down the curtain* a little? That is fine, but will you please **pull it down* a trifle more?" See that the child has taken his hand from the curtain before being instructed to make the second adjustment. Also, be sure that he approaches the shade so that it can be as readily pulled with one hand as with the other.

"Now will you please *take the articles out of the desk* one at a time and place them on the top of the desk?" The first two responses are recorded for this item. The tester picks up the articles and places only the tablet in front of the child, and says, "Please *tear out* one sheet of paper and *turn over* the paper."

The tester then puts the pencil and pencil sharpener in front of the child, and says, "Please **sharpen the pencil.*" The activity is scored by recording the hand that does the actual turning.

The tester next places the pencil directly in front of the child and continues, "*Write* your name at the top of the paper. No, please *erase it*, and *write it* at the bottom of the sheet. Now will you please *draw* a picture; and when you are finished, *put the pencil in the desk.*"

Then a box of crayons is placed directly in front of the child by the tester who says, "Now *color* the picture." Note which hand is used in *taking out each* crayon (score the first two), in *coloring*, in *placing each crayon* in the box (score the first two), and in *closing the box*. In case the child does not voluntarily use more than one color, the tester may say, "Let's use another color for this part of the picture." As soon as the child is through and no longer has his hands on the box, say: "Let's look in the box to see if we have all the crayons in it." This gives a second score for *closing the box*.

Next the scissors are placed directly in front of the child. "Cut off the bottom line, please." Note the hand used in *picking up the scissors*. "Now *put the scissors in the desk.*"

The tester next arranges the blocks in a straight line from one side of the table to the other, the middle block being directly before the child, and says, "*Point to A*; *point to G*," etc. Be sure to see that the child is relaxed and ready to start with neither hand favored before the second letter is named. The first two letters, the only ones scored, are called off in a definite order, the sixth from the left being called first and the third from the left next. "Now pile the blocks one on top of the other to make a pile eight blocks high, being sure to *pick up* only one block at a time." Note the hand used in *picking up*, as well as in *placing*, the second and fourth blocks. The blocks are again put on the desk, and the tester instructs the child by saying, "Please carry these over to that desk one at a time." If there is not a second desk handy, some other surface may be substituted. Record the hand used in *picking up* first and third blocks. Do not record the hand used in carrying the blocks.

Next the tester places five flash cards on the top of the desk and says, "Please *point* to the cards as I name them. Call first for the second from the left and then for the fourth from the left. Now please *pick them up* one at a time and now *lay them down* one at a time." (For these last two items, record the first two responses in each case.)

Now a tablet is placed directly in front of the child, and the tester says, "Please *tear out* another sheet of paper; *turn over the sheet of paper;* and **fold it* in the middle." The hand which actually creases the folded edge is the one recorded.

Next the pencil sharpener and pencil are again presented in the midline, and the child is instructed to **sharpen the pencil.*

Following this activity, the picture book is presented in the middle of the desk, and the tester says, "Please **turn to* the picture of Casper Carrot." Any picture will do so long as two or more pages are turned. Record the hand used in turning each of the first two pages. Then after the pencil is again presented directly in front of the child, the examiner says, "Now *draw* the picture, and when you are through put the pencil in the desk."

Next the book is removed and the scissors are placed directly in front of the child, and the tester says, "Please cut out your picture, and then *put the scissors in the desk.*" Also, the hand used in *picking up* the scissors is recorded.

The bottle of ink and the pen are next presented directly in front of the child, care being taken to see that the pen is placed on neither side of the ink bottle. The tester says, "Please **pick up the pen*, **take off the top*, (record the hand doing the actual turning) put them both down, **take the top off the ink bottle*, **fill the pen*, and **write* your name. Then **put the top back on the pen.*" While the child is doing these things, the tester should put the top back on the ink bottle. The tester next says, "Please *erase* your name. Now **fold* the paper so that the ink spot will not show. **Pick up* the pen, **take off the top*, put down the pen, **take the top off the ink bottle*, **fill the pen*, **write* your name, and **put the top on the pen.*"

The last two activities are concerned with the blackboard. The child is directed to stand squarely in front of it where the chalk can be as easily reached with one hand as with the other. The tester says, "Please **write* your name." Score the hand used in **picking up the chalk* as well as in *writing*. Lastly, the tester asks the child, "Will you please **print* your name?" This gives the second scores for **picking up* the chalk.

This concludes the test as it was revised leaving out the invalid items.†

Any remarks by the child or further directions by the tester, in case there is difficulty in grasping the directions, need not be curtailed, provided they do not influence hand preference.

---

† The items shown by the present study to be invalid were: clasping the hands, raising the hand, putting tablet in desk, picking up pencil, opening crayon box, placing crayon box in desk, cutting with scissors, and erasing the blackboard. In this study, the test was given with these items included—the form of the test here presented is the final revision.

The time required to give the test is twenty to twenty-five minutes per child. No cajoling was necessary to gain the desired response; indeed, the detail of the picture drawn had to be curtailed due to the time element involved.

The tests were given in a manner that made it possible to score them in two ways:

1. The Johnson-Duke method was used, applying their original number of responses to the original twenty-six items.

2. The revised method was used, scoring each of the forty items only twice each.

Thus, two DQ's were obtained for each test and retest, and this made possible a direct comparison of the two methods of testing and scoring. The retest provided a means of measuring the consistency or reliability of both forms of the test.

III. SUBJECTS

The one hundred children tested were all seven-year-olds, ranging in age from seven years, four days to seven years, eleven months, twenty-nine days. The median age was seven years, five months, twenty-seven days. There were fifty boys and fifty girls. All cases of known mental deficiency, speech defects, and physical abnormalities which might interfere with hand usage were excluded.

Sixty-two of the children were tested in the Whittier School, Royal Oak, Michigan, and the remaining thirty-eight in the Ninth Ward and Seventh Ward schools of Eau Claire, Wisconsin. A retest was given within a week to each child, sixty-two per cent receiving a retest the day following the first test.

IV. RESULTS

Reliability.—Each child's DQ for both test and retest (T_1 and T_2 respectively) was computed on the portion of the test used and scored by Johnson and Duke. These Johnson-Duke scores were subjected to correlation and found to yield a test-retest coefficient of $.80 \pm .036$. A further check on the reliability of the test was made as follows: The DQ's of each child for T_1 and T_2 were subtracted, the difference between them so determined, and the standard error of the difference then established. A difference three or more times its standard error is to be regarded as significant. Of the one hundred children tested, ten presented a difference in DQ between T_1 and T_2 which was more than three times its standard error.

The revised scores for T_1 and T_2 yielded a correlation coefficient of .90 ± .019. The difference in DQ for each child between T_1 and T_2 was then computed. It was found that by using the revised method of scoring and the forty items, only four children presented a difference of DQ between T_1 and T_2 which exceeded three times the standard error, and so were statistically significant.

Correlations were made between Johnson-Duke T_1 and revised T_1 scores; they yielded a coefficient of correlation of .89 ± .021. The correlation coefficient between Johnson-Duke T_2 and revised T_2 scores was .89 ± .021. Thus, although the highest correlation coefficient is for the revised T_1 and T_2 scores, both tests are rather highly reliable and agree closely with each other.

Since the revised scores were shown to be more reliable, the reliability of the revised test was further evaluated in another manner. The DQ's of each activity for T_1 and T_2 were found by combining all the responses on each item as performed by all the children. The difference between the T_1 and T_2 scores thus obtained was computed for each activity. Only one item presented a difference 2.7 times as great as its standard error, the rest falling below 2.0, so that the difference between T_1 and T_2 scores for all activities fell below three times the standard error. It is to be concluded that the revised test, considered item by item, is highly reliable.

Validity of the Test Items.—It was necessary, in the further improvement of the test, to determine which items were most highly significant in determining handedness. For this purpose, the upper and lower quartiles of the distribution were compared. The average DQ's on T_1 (revised test) for each quartile for each separate activity were computed, and the difference between the average DQ's of the two quartiles, together with the standard error of the difference, were also computed. It was found that thirty-two of the forty activities were statistically valid, in the sense that they clearly differentiated the lower from the upper quartile of the distribution. These thirty-two items are the ones included in the test as it is described earlier in this paper.

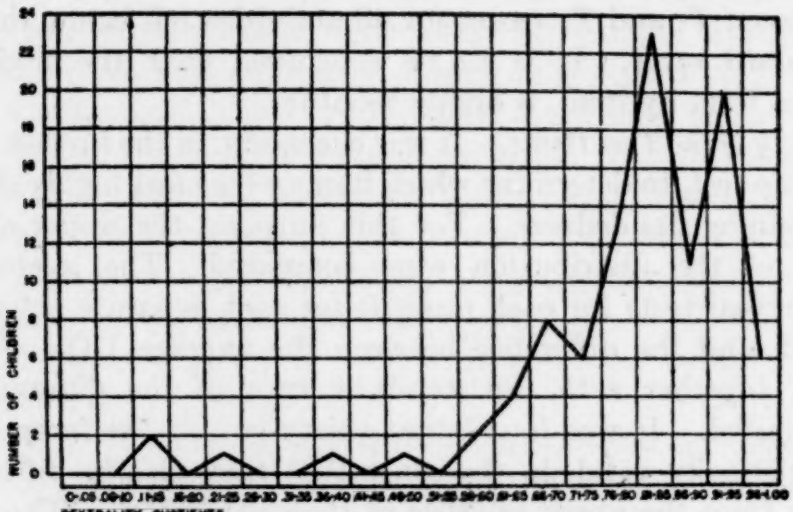
In general, the internal consistency of the test is shown to be relatively high in terms of this criterion.

Reliability of the Test as Further Revised.—The test was rescored by the revised method after deleting the questionable items. The scores on T_1 and T_2 in terms of DQ's were then correlated. The coefficient of correlation was found to be .91 ± .017.

Norms for Seven-year-olds.—Norms are given below in terms of the test as finally revised. The Johnson-Duke norms for six-year-olds are given for purposes of tentative comparison.

Decile	Six-year-olds, DQ's	Seven-year-olds, DQ's
0	.03	.12
10	.44	.63
20	.60	.71
30	.67	.78
40	.73	.81
50	.78	.83
60	.82	.85
70	.85	.88
80	.89	.92
90	.93	.95
100	.99	1.00

The median score for the entire group is .83. The median score for the girls is .85, and the median score for the boys is .82.



DISTRIBUTION OF DEXTRALITY QUOTIENTS

The graph of the distribution of DQ's shows an essentially single mode curve skewed toward the right-handed side of the distribution. Johnson and Duke found essentially the same type of distribution at the six-year level.

Effect of Age on Hand Usage.—In comparing the norms given here for seven-year-olds with the Johnson-Duke norms for six-year-olds, it must be noted that while both sets of norms are based on essentially

the same test, somewhat different forms of the test were used. Also, the present norms are based on measurements of one hundred children, while Johnson and Duke tested only fifty children. With these facts in mind, it appears reasonable to conclude, nevertheless, that the entire distribution of scores shifts toward the right-handed end of the scale when we go from the six-year level to the seven-year level. The median for the six-year-olds is .78. The median for the seven-year-olds is .83. A comparable shift is shown at all the deciles, the greatest shift being in the "left-hand" half of the distribution. These figures suggest that children become more right-handed in terms of hand usage as they become older. The theoretical implications of this finding, and the finding itself, can be better evaluated when other studies, now in progress, are completed.

V. SUMMARY AND CONCLUSIONS

The Iowa Hand Usage Test as originated by Johnson and Duke was administered in a revised form to one hundred seven-year-olds, fifty boys and fifty girls. The sample was random, except that an equal number of each sex was selected, and children with speech defects, known mental deficiency, and physical abnormalities which might interfere with hand usage were excluded. In the course of the investigation an improvement of the test was achieved by increasing the number of items, excluding invalid items, and by scoring only two responses per item. The test is scored in terms of the dextrality quotient, DQ, which is to be defined as the percentage of achievement in any test of handedness which is to be credited to the right hand.

Results of the investigation may be summarized as follows: Hand usage DQ norms for seven-year-olds, in terms of the Iowa Hand Usage Test as here revised, are made available in this report. A graph of the distribution indicates that hand usage DQ's are not distributed in a clearly bimodal fashion, but rather in a skewed and essentially unimodal manner. The median DQ for the one hundred seven-year-olds is .83. This means that in terms of the test here used the median seven-year-old is eighty-three per cent right-handed as far as hand usage is concerned. It is shown that on the average girls exhibit a slightly higher degree of right-hand usage at seven years than do boys. Comparison of the seven-year-old norms with the tentative six-year-old norms of Johnson and Duke indicates that in terms of hand usage, the degree of right-handedness increases with age.

It is of particular interest that among the seven-year-olds, as among Johnson and Duke's six-year-olds, there is what might be termed a "consistent inconsistency" in hand usage for the group as a whole. That is, the test-retest correlation of .91, the median DQ of .83, and the reliability and validity of each test item indicate that in general the seven-year-old does not use the left hand now and then in a random manner, but that he uses it to a certain quite definite degree and for quite well designated activities. He is practically one hundred per cent right-handed for certain activities, and practically one hundred per cent left-handed for others. In terms of the test used, he does not present to a notable degree the type of ambidexterity characterized by inconsistency of hand usage with respect to any given activity.

BIBLIOGRAPHY

1. Johnson, Wendell and Duke, Darlene: "The dextrality quotients of fifty six-year-olds with regard to hand usage." *J. Educ. Psychol.*, Vol. XXVII, 1936, pp. 26–36.

A COMPARATIVE STUDY OF THE INTELLIGENCE OF DELINQUENT AND DEPENDENT BOYS

JOSEPH E. MOORE

George Peabody College for Teachers

The close relationship between delinquency and dependency has been consistently shown by sociological and economic surveys. However, few studies on the psychological relationship between the two groups, as measured by objective tests, have been made. Investigations on the intelligence of delinquents have been almost entirely based on comparisons with normal groups rather than with dependent groups of approximately the same social and economic background.

The intelligence of both delinquent and dependent children has generally been found to be below the test norms. In the few previous studies in which delinquents and dependents have been compared in intelligence, the number of subjects used has been inadequate, or various selective factors have not been controlled, and the results are rather contradictory. Pintner[11] has perhaps made the most thorough comparative study and summarization of the work done on the intelligence of delinquents and dependents as related groups. In his comparative study his subjects consisted of one hundred juvenile court delinquents of both sexes, and forty-eight dependent children. Of the delinquent group forty-three per cent were more than three years retarded according to scores on the Binet-Goddard intelligence test, and twenty-three per cent, or approximately one-half as many dependents, showed the same amount of retardation. The Binet-Goddard test, however, was not standardized at the time this study was made, as Pintner points out, and it has since been found too difficult at the upper age levels; the subjects included both sexes; the delinquents were not institutional cases although the dependents were; and the number of dependent children was quite small. A second study in which the intelligence of delinquents and dependents is compared is that of Horn,[8] who compared the intelligence of eighteen dependents with that of twenty-four delinquents of both sexes. Of the delinquent group eight of the subjects, or thirty-three per cent, were classed as feebleminded, and seven or thirty-nine per cent of the dependents were classed as feebleminded. The number of subjects, however, is quite small and the delinquent group included adults as well as younger subjects. A third study is that of Slawson[14] who made a rather general comparison of the intelligence of his groups

of incarcerated delinquent boys in New York with the group of dependents (both sexes) in Connecticut county homes, as reported by Cobb.[5] The Binet test was used in both studies. Slawson found that the delinquent boys performed even better than the dependents in Cobb's study. However, in many training schools the population is somewhat a select group, since in many States there are intelligence testing programs in the juvenile court clinics, and delinquents testing definitely feebleminded are committed to other institutions than training schools. The State of New York, in fact, has a separate institution for feebleminded delinquents.[13] In testing incarcerated juvenile delinquents, then, it should be determined beforehand and definitely stated whether or not such a selective process has taken place in the group studied in order to avoid confusion or contradictory results when comparisons are made with other investigations.

In the present study two groups, one of delinquent boys from the Tennessee State Training and Agricultural School and the other of dependent boys from the Tennessee State Industrial School, were used as subjects. The Industrial School, a State orphanage for children who have become public charges through inability of private agencies to provide for their welfare, does not accept any children who have been charged with or convicted of a violation equal to the civil or criminal laws of the State.[1] The Training School, on the other hand, receives only the delinquent boys convicted of criminal offense.[1] Neither group had been selected previously on the basis of intelligence test results, since the State of Tennessee at the time the present study was made had no provision for the mental testing of delinquents or dependents.

The delinquent group was composed of one hundred fifty-two boys and the dependent group of one hundred fifty-seven boys. Both groups of subjects consisted of white, native-born boys of native-

TABLE I.—NUMBER OF DELINQUENT AND DEPENDENT BOYS IN EACH CHRONOLOGICAL AGE GROUP FROM FOURTEEN THROUGH EIGHTEEN YEARS

Group	Fourteen years	Fifteen years	Sixteen years	Seventeen years	Eighteen years	Total
Delinquent...	28	27	37	27	33	152
Dependent....	47	41	35	18	16	157

born parents. The chronological age-range of the subjects in both groups was fourteen through eighteen years, inclusive, age being

reckoned from the nearest birthday. Table I shows the number of subjects from each school in the different age groups.

The median chronological age of the delinquents was sixteen years, three months, and of the dependents, fifteen years, five months.

Selection of subjects was made by age-grade sampling. The school grade range of the delinquent group was from the fourth through the eleventh, inclusive, and of the dependents from the fifth through the tenth grades, inclusive. Since there was a much larger number of delinquent than dependent boys in the primary grades, it should be mentioned that the delinquent boys were somewhat the favored group in this study.

All tests were given during May and June, 1935, by the author.

The tests used to measure intelligence were the Otis Self-Administering Test, Intermediate and Advanced Forms, and the Myers Mental Measure.

In the scoring of Otis IQ two methods were used. The IQ's were first calculated on the basal age level used in the Otis scoring key,

TABLE II.—MEASURES OF CENTRAL TENDENCY AND VARIABILITY IN OTIS IQ OF ONE HUNDRED FIFTY-TWO DELINQUENT BOYS AND ONE HUNDRED FIFTY-SEVEN DEPENDENT BOYS

Measure of central tendency and variability	One hundred fifty-two delinquent boys		One hundred fifty-seven dependent boys	
	IQ with Otis norm	IQ with fourteen year norm	IQ with Otis norm	IQ with fourteen year norm
Median	68.9	74.4	86.4	89.9
Q_1	59.4	64.5	75.4	80.8
Q_3	79.1	83.2	94.3	101.7
Q	9.9	9.4	9.5	10.5
Mean	70.1	75.5	84.2	89.0
SD	15.9	16.5	13.8	14.5
Standard error of mean	.91	.95	.78	.82

and then recalculated on a fourteen-year basal age level, since many investigators, especially those who have tested delinquents, tend to consider sixteen years rather high as a basal age for adult intelligence; among these are Healy and Bronner,[7] Doll,[6] Boynton,[4] and Slawson.[14] Slawson preferred to use both methods for the calculation of IQ in giving the Binet test to groups of delinquents. The lower basal age

was also used in the present study in order to compare more effectively the results of the Otis and Myers intelligence tests, both of which were given, since the Myers Mental Measure is not standardized above the fourteenth year.

Table II shows the results of the Otis test on the delinquent and dependent groups, as expressed in measurements of variability and central tendency.

From Table II the dependents appear to exceed the delinquents approximately sixteen points in median IQ at both basal age levels, although the Standard Deviation is greater for the delinquents. That the difference is reliable is shown by the fact that the Critical Ratio* between the Mean IQ of the two groups is 7.8. Hence, it can be rather definitely stated that on the Otis test the group of dependent boys is superior to the group of delinquent boys, although

TABLE III.—MEDIAN OTIS IQ AND MEDIAN YEARS OF SCHOOL RETARDATION OF ONE HUNDRED FIFTY-TWO DELINQUENT BOYS AND ONE HUNDRED FIFTY-SEVEN DEPENDENT BOYS AT EACH CHRONOLOGICAL AGE LEVEL FROM FOURTEEN THROUGH EIGHTEEN YEARS

Age	One hundred fifty-two delinquent boys				One hundred fifty-seven dependent boys			
	No.	IQ Otis norm	IQ fourteen year norm	School years retarded	No.	IQ Otis norm	IQ fourteen year norm	School years retarded
14	28	73.4	76.0	2.8	47	85.5	86.5	.9
15	27	63.8	68.2	3.3	41	86.8	90.8	2.1
16	37	66.3	73.8	4.5	35	87.1	93.5	1.6
17	27	69.2	76.9	5.0	18	91.9	100.3	1.5
18	33	70.5	75.6	5.5	16	82.5	89.5	2.8
Total group..	152	68.9	74.4	4.5	157	86.4	89.9	1.6

both delinquents and dependents are below normal, when judged by either basal age norm, the delinquents showing approximately twice the amount of retardation that the dependents show. The difference in median IQ found when the fourteen-year basal age norm is used

* Calculated by the formula:

$$CR = \frac{d}{\sigma_d}$$

and when the Otis norm is used is similar to that found by Slawson[14] using the fourteen- and sixteen-year norms in determining Binet IQ on a group of delinquent boys of approximately the same chronological age as the group of delinquents in the present study.

When the median Otis IQ's and median years of school retardation of the two groups are compared at each chronological age level, the difference between the two groups remains consistent with the data presented in Table II as to the median IQ, and the delinquents show a much greater amount of school retardation than do the dependents.

From the above table it appears that the delinquent group is consistently more retarded in both IQ and actual school-grade achievement than the dependent group at each chronological age level

TABLE IV.—DISTRIBUTION, PERCENTAGE, AND FREQUENCY OF OTIS IQ OF ONE HUNDRED FIFTY-TWO DELINQUENT BOYS AND ONE HUNDRED FIFTY-SEVEN DEPENDENT BOYS WHEN BOTH OTIS NORM AND FOURTEEN YEAR NORM ARE USED TO DETERMINE IQ

IQ interval	One hundred fifty-two delinquent boys				One hundred fifty-seven dependent boys			
	Otis norm		Fourteen year norm		Otis norm		Fourteen year norm	
	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent
120–above.....	1	.66	2	1.32	...	.....	3	1.91
110–119.....	3	1.97	6	3.94	6	3.82	8	5.10
100–109.....	4	2.63	8	5.26	14	8.91	25	15.92
90– 99.....	12	7.90	11	7.24	35	22.29	42	26.75
80– 89.....	15	9.87	20	13.16	44	28.03	32	26.75
70– 79.....	33	21.71	42	27.63	39	24.84	21	13.38
60– 69.....	43	28.29	42	27.63	11	7.01	11	7.01
50– 59.....	41	26.97	21	13.82	8	5.10	5	3.18
Totals.....	152	100.00	152	100.00	157	100.00	157	100.00

included in the present study. In the dependent group IQ tends to increase somewhat with chronological age until the eighteenth year, although the numbers involved are too small for any general conclusion. However, a similar trend is seen in the delinquent group after the fourteenth year, the fourteen-year-old group being the exception. School retardation, also, tends to increase slightly with age in the

delinquent group, although there is no definite trend in the dependent group as to median school retardation.

Table IV gives the distribution and frequency of the Otis IQ of the delinquent and dependent boys at both basal age levels.

From Table IV it will be seen that although the dependents present a somewhat normal distribution curve which falls approximately one class interval below that of Terman's group of unselected school children, the delinquent boys present a distribution curve skewed sharply toward the higher IQ levels. The fact that the Otis test does not score below fifty IQ, as well as the method of selection, may account in part for the skewness of the distribution curve for the delinquents, although all subjects tested made at least one correct response on the Otis intelligence test.

More than fifty per cent of the delinquents score below seventy IQ when the Otis norm is used, and more than forty per cent score below seventy when the fourteen-year norm is used. The percentage of dependents scoring below seventy IQ is much smaller than that of the delinquents, only ten to twelve per cent of this group scoring below seventy IQ. On the other hand, only approximately thirteen per cent of the delinquents attain a score of ninety IQ or above when the Otis norm is used, and approximately thirty-five per cent of the dependents attain this score with the Otis norm. When the fourteen-year basal age is used approximately eighteen per cent of the delinquents and forty-seven per cent of the dependents score ninety IQ or above.

Although no diagnosis of feeblemindedness is made in this study, it should be pointed out that since the accepted estimation is that only two per cent of the general population would test below seventy IQ, there seems to be a preponderance of low intelligence in both the delinquent and dependent groups, the percentage of dependents testing below seventy being five times as great as that considered normal, and the percentage of delinquents testing below seventy IQ being twenty times as great as the accepted norm, when the lower basal age level is used to determine IQ. When the Otis norm is used, the percentage below seventy IQ is increased in both the dependent and the delinquent groups.

When the results of the Otis test in the present study are compared with the findings in other investigations on groups which had not been selected previously on the basis of intelligence testing, and especially on groups of delinquent boys of approximately the same geographical area, the results are quite similar. Boynton[3] reports a median Binet

IQ of sixty-nine for a small group of twelve-year-old delinquent boys at the Kentucky Houses of Reform. Knight[2] found a median Binet IQ of 73.9 for one hundred forty-seven delinquent boys in a North Carolina training school, and Moorer,[9] using the Henmon-Nelson test, reports a median IQ of 69.7 for a group of two hundred seventy delinquent boys at the Tennessee Training and Agricultural School. Moorer's investigation was made one year previously to the present study and probably included a number of subjects tested in the present study. Table V gives a frequency distribution of IQ in terms of percentage comparing Moorer's and Knight's groups of delinquent boys with the group in the present study.

TABLE V.—FREQUENCY DISTRIBUTION OF IQ IN TERMS OF PERCENTAGE OF THREE GROUPS OF DELINQUENT BOYS

Frequency interval of IQ	Otis group one hundred fifty-two boys		Henmon-Nelson two hundred twenty-four boys	Stanford-Binet one hundred forty-seven boys
	Otis norm	Fourteen year norm	Sixteen year norm	Sixteen year norm
110–above	2.63	5.26	1.34	1.2
100–109	2.63	5.26	2.23	3.6
90– 99	7.90	7.24	6.25	8.2
80– 89	9.87	13.16	8.93	15.3
70– 79	21.71	27.63	23.66	24.2
60– 69	28.29	27.63	29.02	33.6
50– 59	26.97	13.82	20.98	12.4
40– 49	.....	.....	7.59	1.5
Total	100.00	100.00	100.00	100.00

The two hundred twenty-four subjects in the Henmon-Nelson group consist only of the boys in Moorer's study within the age group included in the present study, those subjects below fourteen years and above eighteen years being omitted from Table V and the frequency distribution recalculated for this group in order to make a more accurate comparison with the group in the present study. The age range of the Binet group was from eight to twenty-one years, inclusive, the median chronological age being fifteen years, ten months, or approximately five months below the median age of the Otis group.

From Table V it appears that the skewness in distribution of IQ shown by the Otis test results is also present in the Henmon-Nelson group of delinquents at the same institution, and in the Stanford-Binet group. The frequency distribution of the Stanford-Binet IQ in Table V, however, seems to agree more closely with the Otis findings when the fourteen-year norm is used to determine IQ. The similarity is not lacking, however, even when the fourteen-year norm is omitted in comparing the three studies.

The percentage of low-grade intelligence appears to be rather high, then, or approximately fifty per cent, not only in the present study but also in other studies of incarcerated delinquent boys who have not been selected on the basis of intelligence testing in juvenile court or elsewhere before entering the training school, and who are from approximately the same geographical area as the subjects in this study.

FINDINGS ON THE MYERS MENTAL MEASURE

The Myers Mental Measure, a non-verbal intelligence test, was given to one hundred forty-eight of the delinquents and one hundred eighteen of the dependents who had taken the Otis test. The correlation between the Myers raw scores and the Otis IQ was $44 \pm .06$ for the dependent boys and $50 \pm .04$ for the delinquent group.

The median chronological age of the delinquents in the Myers group was sixteen years, five months and that of the dependents fifteen years, eleven months. The age-range was the same as that for the entire group taking the Otis test, being fourteen through eighteen years, inclusive. The median chronological age of the delinquents in the Myers group was sixteen years, five months, and that of the dependents was fifteen years, eleven months. The median school grade of the delinquents taking the Myers test was 5.25 and of the dependents was 8.29. The median Otis IQ for this group of delinquents was 74.0 when calculated by the fourteen-year norm, and for the dependents taking the Myers was 89.9. The fourteen-year basal age was used in computing the Otis IQ's of both groups of subjects taking the Myers Mental Measures to insure a more accurate comparison of the test results, since the Myers norm shows no improvement in response after the fourteenth year.

In the following table the raw scores of the Myers test for the delinquent and dependent boys are given according to chronological-age group with the Myers and Otis IQ's for both groups of subjects.

All figures are in terms of median. The Myers IQ's are given only for very general comparison with those of the Otis, the raw scores of the Myers being much more accurate.

TABLE VI.—MYERS RAW SCORES, MYERS IQ'S, AND OTIS IQ'S OF ONE HUNDRED FORTY-EIGHT DELINQUENT BOYS AND ONE HUNDRED EIGHTEEN DEPENDENT BOYS AT EACH CHRONOLOGICAL AGE LEVEL FROM FOURTEEN THROUGH EIGHTEEN YEARS

Age	One hundred forty-eight delinquent boys				One hundred eighteen dependent boys			
	No.	Myers raw score	Myers IQ	Otis IQ	No.	Myers raw score	Myers IQ	Otis IQ
14	25	41.9	85.7	75.5	30	50.0	100.0	86.0
15	27	54.4	100.0	66.0	33	47.8	92.9	90.0
16	37	40.9	78.5	73.3	29	51.9	100.0	96.0
17	26	45.0	85.0	76.4	13	56.3	100.0	100.3
18	33	46.3	92.9	74.0	13	49.2	100.0	89.5
Group medians...	..	42.6	85.7	74.0	..	50.0	100.0	89.9

From Table VI it would appear that both the delinquent and dependent groups tend to score higher, or nearer normal, on the Myers Mental Measures than on the Otis intelligence test, even when the fourteen-year norm is used to determine Otis IQ. Neither is the difference between the dependent and delinquent boys as great when judged by the Myers results as when judged by the Otis, the Critical Ratio of the difference between the two groups of subjects being 3.6 for the Myers raw scores and 7.8 for the Otis IQ's. The dependent group according to the median score on the Myers Mental Measures is normal. The median score of the dependent group according to Table VI reaches the adult norm for the Myers Mental Measures. The delinquent group exceeds the norm only at the fourteen-year level, the median score for the group being 7.4 points below normal and approximately fourteen points in Myers IQ below the norm. However, it is inadvisable to draw any definite conclusions regarding the showing of the delinquent or the dependent group on the Myers test, since Myers' findings[10] show a reliable difference between the scores of children of low economic origin and those of average and high economic origin in the same school. However, in-so-far as the delinquents and the dependents in the present study

may be considered as belonging to approximately the same social group, it can be rather definitely stated that the dependents exceed the delinquents in attainment on the Myers Mental Measures. On the other hand, it may be questionable to assume that the delinquent group in this study is definitely below normal on the Myers, for their social group, although the delinquents are quite definitely below the norm set by unselected school children, as well as below the score of the control group in this study.

The following table gives the frequency distribution and percentage of Myers raw scores of one hundred forty-eight delinquent boys and one hundred eighteen dependent boys.

TABLE VII.—FREQUENCY DISTRIBUTION AND PERCENTAGE OF MYERS RAW SCORES OF ONE HUNDRED FORTY-EIGHT DELINQUENT BOYS AND ONE HUNDRED EIGHTEEN DEPENDENT BOYS

Frequency interval	One hundred forty-eight delinquent boys		One hundred eighteen dependent boys	
	Number	Per cent	Number	Per cent
80–89	1	.68	1	.85
70–79	3	2.03	7	5.93
60–69	17	11.48	14	11.86
50–59	22	14.80	37	31.36
40–49	44	29.73	31	26.27
30–39	38	25.68	20	16.95
20–29	20	13.51	8	6.78
10–19	3	2.03		
Total.....	148	100.00	118	100.00

The Myers raw scores of both groups of subjects in the above table tend to follow a somewhat normal distribution curve, although the scores of the delinquent group fall approximately one-class interval below those of the dependents, the largest percentage of the delinquents falling in the 30–39, 40–49 score grouping and the largest percentage of the dependents falling in the 40–49, 50–59 grouping.

SUMMARY OF FINDINGS

Both the delinquent boys and the control group of dependent boys in this study were found to be rather definitely below the norm on the Otis S-A intelligence test. When the basal age of fourteen years

was used to determine IQ, the median IQ of both the delinquent and dependent groups was found to be slightly higher, although it remained rather definitely below the accepted norm for unselected school children for both groups of subjects.

When the median Otis IQ's of the delinquent boys and the dependent boys in this study were compared, the dependents were found to have a definitely higher median IQ than the delinquents, the difference being statistically reliable.

The results found in giving the Otis test to the delinquent group in this study agree rather closely with those in other studies of the intelligence of delinquent boys in which the group of subjects was not selected previously on the basis of intelligence, and were of approximately the same geographical area.

On the Myers Mental Measures, a non-verbal test, the dependent group in this study were found to score barely normal. The delinquent group was found to score approximately two years below normal for their chronological age. The difference between the delinquent and dependent groups in attainment on the Myers Mental Measures is statistically reliable, although the difference between the two groups in the Myers scores is less than that in the Otis IQ's.

REFERENCES

1. *Biennial Report* of the Department of Institutions of Tennessee, 1930–1932, pp. 225–231; 232–259.
2. *Biennial Report* of the North Carolina Charitable, Penal, and Correctional Institutions, 1930–1932, pp. 283–294.
3. Boynton, P. L.: "A Study of the Mental Development of Twelve-Year-Old Boys in the Kentucky Houses of Reform." *J. Delinq.*, Vol. X, 1926, pp. 532–534.
4. Boynton, P. L.: *Intelligence: Its Manifestations and Measurement.* New York: D. Appleton & Co., 1933.
5. Cobb, Margaret E.: "The Mentality of Dependent Children." *J. Delinq.* Vol. VII, 1922, pp. 132–140.
6. Doll, E. A.: "The Average Mental Age of Adults." *J. Appl. Psychol.*, Vol. III, 1919, pp. 317–328.
7. Healy, W. and Bronner, A. F.: *Delinquents and Criminals: Their Making and Unmaking.* New York: Macmillan Co., 1926.
8. Horn, J. L.: "A Study of Fifty-Three Juvenile Court Wards." *J. Delinq.*, Vol. III, 1918, pp. 189–209.
9. Moorer, Sam H.: *A Reformatory as an Educational Institution.* Master's Thesis (unpublished). Vanderbilt University, Nashville, Tenn., 1935.
10. Myers, G. C. and Myers, C. E.: "Comparative Intelligence Ratings of Three Social Groups within the Same School." *School & Soc.*, Vol. XIII, 1921, pp. 536–539.

11. Pintner, Rudolf: "One Hundred Juvenile Delinquents Tested by the Binet Scale." *Ped. Sem.*, Vol. XXI, 1914, pp. 523–531.
12. Pintner, Rudolf: *Intelligence Testing.* New York: Henry Holt & Co., 1931, Ch. XVI, "The Delinquent," and Ch. XVII, "The Dependent."
13. Robinson, L. N.: "Institutions for Defective Delinquents." *J. Crimin. Law and Criminol.*, Vol. XXIV, 1933, pp. 352–399.
14. Slawson, John: *The Delinquent Boy.* Boston: Gorham Press, 1926.
15. Traxler, A. E.: "Reliability and Validity of the Otis IQ." *J. Appl. Psychol.*, Vol. XVIII, 1934, pp. 241–251.

ON CORRELATIONS BETWEEN MEASUREMENTS WHICH ARE NOT NORMALLY DISTRIBUTED

EDWARD L. THORNDIKE

Institute of Educational Research, Teachers College, Columbia University

The measurement of the interrelation of two series of measures, when one or both are skewed or otherwise not normally distributed, is difficult. In some such cases it is reasonable to avoid the difficulty by shrinking or expanding the scale until a normal distribution is obtained or by using the order of magnitude.[1] But when the units of the scale are minutes, pounds, persons, dollars, or other presumably equal amounts, it is unsound to change the measures for the sake of convenience in computing correlations.

A method which has two advantages is to express the measures of a series as deviations from its median, and measure the intercorrelation by the percentage of like-signed pairs. This is very economical of time; and the percentage of like-signed pairs expresses the degree of intercorrelation more clearly and conveniently than the sum of the differences in rank or than the sum of the squares of the differences in rank. A coefficient computed from it by Sheppard's formula [$r = \cos \pi(100 - L)$] is at least as "true" and useful as a coefficient computed from the differences in ranks by either of Spearman's formulas. The method has the disadvantage of utilizing the data less fully than the product-moment method.

It is of some interest to observe the results obtained when r is computed for irregularly distributed series by the product-moment method $r = \Sigma xy/\sqrt{\Sigma x^2 \cdot \Sigma y^2}$. I shall take x and y as deviations from the median, instead of from the mean, for reasons which will appear in what follows. I shall use as variables to be correlated various series of measures from each of one hundred seventeen cities, the units of the scales being a person, a dollar, a death, an acre, or a ratio (*e.g.* that which the number of persons of one sort is to the number of another sort, or that which the number of dollars spent in one way is to the number spent in another). Some of the distributions are somewhat skewed, others are greatly skewed, others are approximately normal. The skewness is commonly toward the high end of

[1] The formulae for correlating by using the order of magnitude do not use only it, but also some assumption about the spacing within the order (commonly the assumption that the spacing is as it would be in a normal distribution). They thus actually alter the measures from their face value.

the scale, but there are a few cases of negative skewness. Sample distributions are shown in Table I.

Let us first compare the product-moment coefficients with those computed from the percentage of like-signed pairs by

$$r = \cos \pi(1.00 - L).$$

In fifty-seven correlations where neither of the two distributions departs notably from normality, the Pearson coefficients, as we should expect, do not depart appreciably from the like-signed pair coefficients. The latter run about five per cent higher, but this may be due in part to my arbitrary treatment of pairs involving one or two zeros, $\frac{\left(\frac{L}{L+U}+\frac{1}{2}\right)}{2}$ of which were treated as L. In part it is due to the fact that even these distributions are not perfectly normal.

When, however, one distribution is normal and the other moderately skewed, the product-moment coefficients are clearly lower than the others, the median of seventy-nine cases being .58 for $\cos \pi(1.00 - L)$ to .51 for $\Sigma xy/\sqrt{\Sigma x^2 \cdot \Sigma y^2}$.

When both distributions are skewed in the same direction, the same holds good, the median of forty-seven such cases being .66 to .59.

When the two distributions are skewed in opposite directions, or when one distribution approaches normality and the other is much skewed, the difference is much larger, the median reduction for sixty-four such cases being from .53 to .38.

It cannot be assumed that the $\cos \pi(1.00 - L)$ coefficients are "true" measures of the closeness of the interdependence of the variables concerned, and that the tendency of the product-moment coefficients to run lower is a constant "error." But the former do have the advantage of being less influenced by the status of measures in or near the tails of the distributions where the irregularities are greatest. And the latter do involve an error in the sense that the sum of the xy products will be less than the geometric mean of the Σx^2 and Σy^2 even though the correlation is perfect to the extent that the rank orders are identical, and the sum of the xy products is as great as it can be from two such distributions. The former would then (supposing that my arbitrary treatment of pairs involving zero measures had been the wisest possible) be nearer the measures of closeness of interdependence which full study of each relation-plot would suggest. Of course no well-advised person would use Pearson

coefficients for such material as ours, but he might be afraid to use the others. The fact that they rise above the Pearson coefficients in the way and to the extent that they do is in their favor.

TABLE I

1 = the number of Negroes (per 100,000 population).
2 = the number of acres of public parks (per 100,000 population).
3 = the value of public property used as schools, libraries, museums, asylums, hospitals and parks (per 100,000 population).
4 = the quotient of the amount spent by certain religious denominations for outside benevolences, divided by the amount spent by them for maintaining their local churches.
5 = a weighted composite of 22 items.
6 = infrequency of reported arrests for homicide (per 100,000 population).
7 = infrequency of reported arrests for rape (per 100,000 population).

	1	2	3	4	5	6	7
−17							1
−16							1
−15							
−14							
−13							1
−12							1
−11							
−10							2
−9					2	3	2
−8					2		3
−7					2	5	1
−6					4	2	4
−5				1	4	1	2
−4		18		4	9	9	3
−3		11	3	8	4	12	10
−2	34	15	14	19	8	7	4
−1	18	9	28	16	16	10	11
0	14	10	26	11	13	16	16
+1	10	6	19	16	18	16	15
+2	6	7	11	18	16	11	15
+3	2	8	7	5	8	16	25
+4	2	5	1	7	4	2	
+5	2	1	3	4	3		
6	2	4	2	3	2		
7	2	3	1	2	1		
8				1			
9	1	2	2				
10	1	2		2	1		
11	1	3					
12	1	2					
13	2	2					
14	2						
15							
16		2					
17		1					
18							
19	1	1					
20		1					
21							
22		1					
23							
24							
25		1					
26		1					
		and 1 at +34					

If we are seeking measures of the closeness of correspondence of non-normally distributed series whose units of measure should be

left undisturbed, in order to compare different pairs of series in respect of closeness, or to discover factors accounting for all the interrelations of the variables, or to build a regression equation to predict some fact from the use of all or many of the variables, or for other purposes for which measures of closeness of correspondence are useful, we may profitably pursue one of four courses: (1) We may plot each relation, study it by refined methods chosen *ad hoc* and give it a coefficient expressing its general closeness of correspondence in some way that will enable us to use it with the others with a minimum of error. This is extremely laborious. (2) We may compute the differences in ranks and use them in formulae adapted more or less well to the forms of the two distributions in question. Even with only five formulae (for symmetrical with moderate skew, symmetrical with much skew, moderate skew with much skew in the same direction, moderate skew with moderate skew in the opposite direction, moderate skew with much skew in the opposite direction) this method should utilize the facts rather fully without excessive labor and without vicious constant errors. This method is thus promising. (3) We may use the sum of the xy products, dividing it by the maximal sum of the xy products which the two series could give. This uses all of every measure just as it is, but is more laborious than the use of differences in ranks. (4) We may use the percentage of like-signed pairs (x's and y's being taken from the median) either as they stand with recognition of their limitations or in formulae adapted to the forms of the two distributions in question (or even, without great error, in Sheppard's formula). This method is of course extremely economical of time.

Any one of these four courses seems to me much better than to profess inability because the data are too irregular.

THE EFFECTS OF EMPHASIZING "HOW TO LEARN" UPON KNOWLEDGE OF COURSE CONTENT AND SCHOOL MARKS

R. W. EDMISTON

Miami University

The purpose of this study was to determine whether or not an emphasis upon how-to-learn would effect the amount of learning (1) within the course where applied, and (2) in other school work. School marks and standardized comprehensive tests or examinations in psychology were used for evaluations. Some differences in comparisons are necessitated in the two experiments due to inability to obtain equated groups in the second experiment.

In the data from experimental results, interesting variations of individuals are generally noticeable. Since teaching deals with individuals, significant information from students' individual records will be presented. The group data offer central tendencies. Individuals' data are given to portray the existing deviations in a truly meaningful manner.

The groups referred to in this experiment were made up of college freshmen. All were registered in the School of Education. At least three-fourths were girls.

PROCEDURES AND DATA

In all control groups the text used was the same.[1] The units followed in general the Parts and Chapters of this book with identical outside references. In both control and experimental groups the introduction developed the idea of psychology as human behavior and learning for personality development. Text and references were the same for the experimental group, but the order was changed to start with the necessary steps of learning after the introductory unit. Each unit of the subject was then used to stress the steps of learning. In order to further this outcome for the student, a perspective, a brief, examples showing relation of information to recalled situations, and final organization were required for each unit. The results of different purposes or motivations were demonstrated. A knowledge of how to answer different types of test questions and the importance of following directions were impressed by analyzing with students their own test errors and variations. Reading

[1] Robinson, E. S.: *Practical Psychology.* New York: The Macmillan Co., 1932.

TABLE I.—POINT AVERAGES IN PSYCHOLOGY OF GROUPS EQUATED ACCORDING TO INTELLIGENCE AND ACHIEVEMENT UNDER DIFFERENT INSTRUCTORS

	Means and standard deviations at end of marking periods with the statistical significance of their differences					
	Five weeks		Eighteen weeks		$\frac{\text{Diff. } m_2 - m_1}{\text{PE diff.}}$	Significance
	m_1	σ	m_2	σ		
Experimental.....	2.22	.55	2.48	.72	3.00	98
Control.....	2.36	.65	2.37	.78	.104	53
$\frac{\text{Diff. } me - mc}{\text{PE diff.}}$.....	−1.89		.109			
Significance.....	90		77			

INTERPRETATION OF TABLE: Reading from left to right after the word, *experimental*, the experimental group's mean point average (m_1) was 2.22 with standard deviation of .55 at the end of five weeks. This group's mean (m_2) was 2.48 with a standard deviation of .72 at the end of eighteen weeks. The difference of the means ($m_2 - m_1$) divided by the probable error of their difference is 3.00. This figure designates that there are ninety-eight chances in one hundred that the difference is a true difference. Similar data appear from left to right after the word, *control*, for the control group. In this case the difference of the control group's means at the end of the eighteen and five weeks' periods divided by the probable error of this difference gives .104 which signifies fifty-three chances in one hundred that the difference is a true difference. Therefore, the experimental group's gain is much nearer a certainty.

The table should also be considered from top to bottom. Under the *five weeks* heading, the mean of the experimental group, 2.22, has the mean of the control group, 2.36, subtracted from it and then divided by the probable error of this difference to give −1.89. This figure signifies that there are ninety in one hundred chances of a true difference in favor of the control group. In a like manner, under the *eighteen weeks* heading, like statistical procedure shows seventy-seven in one hundred chances that a true difference exists which is now in favor of the experimental group. Therefore, the experimental group, showing ninety chances in one hundred of lesser knowledge at the end of five weeks, has improved by the end of eighteen weeks to an extent designated by seventy in one hundred chances of true superiority over the control group.

comprehension and eye movements were tested, and remedial work was given in a limited number of cases. What and how to memorize were stressed. Inductive and deductive reasoning were practiced, and some adjustments of time and place of study were made.

TABLE II.—POINT AVERAGES IN OTHER THAN PSYCHOLOGY COURSES OF GROUPS EQUATED ACCORDING TO INTELLIGENCE AND ACHIEVEMENT

$N = 50$	Means and standard deviations at end of marking periods with the statistical significance of their differences									
	Five weeks		Eighteen weeks		Diff. $m_2 - m_1$ / PE diff.	Significance	Thirty-six weeks		Diff. $m_3 - m_1$ / PE diff.	Significance
	m_1	σ	m_2	σ			m_3	σ		
Experimental	2.23	.52	2.47	.62	3.16	98	2.46	.72	2.77	97
Control	2.32	.69	2.34	.84	.192	55	2.43	.84	1.07	76
Diff. $me - mc$ / PE diff.	−1.08		1.30		.....	..	.28			
Significance	77		81		.....	..	58			

INTERPRETATION OF TABLE: Table II is to be read as described under Table I. Means and standard deviations for the experimental and control groups at the end of the thirty-six weeks' period are added. These additions portray differences or gains between the thirty-six weeks' and the five weeks' periods as designated by $m_3 - m_1$. The final figures in the experimental and control horizontal rows designate that the experimental group's improvement between the five and thirty-six weeks has ninety-seven in one hundred chances of being a true difference while that of the control group has seventy-six in one hundred chances. Under the *thirty-six weeks* column the significant figure, fifty-eight, shows a little indication of superiority of the experimental group but a lessening of this indication of superiority since the end of the eighteen-week period.

First Experiment.—In the first study, control and experimental groups of fifty each were equated by both intelligence and achievement test scores. More than one hundred students were in the experimental group. Only fifty could be equated but this group's intelligence and achievement scores provide curves approximating very closely the normal curve. Comparisons of the point averages[1] in psychology

[1] $A = 4$ points, $B = 3$ points, $C = 2$ points, $D = 1$ point, E and $F = 0$.

at the end of the five weeks' period and at the end of the first semester, when experimentation ceased, are reported in Table I. These data designate the superior improvement of the experimental group.

Comparisons of other than psychology-point averages for the same group for the five weeks, eighteen weeks, and thirty-six weeks are presented in Table II. The greater improvement of the experimental group is again apparent at the end of the eighteen-week period. However, during the second semester, the improvement ceases with the discontinuation of the how-to-study instruction. The control group's improvement continues during this second semester but does not attain that of the experimental group.

The same fifty people did not pursue the psychology course throughout the year. Therefore, Table III does not compare the identical experimental group of Tables I and II with the standards, but a very similar group since the eight persons dropped from the original group were replaced by an equated group. The standards were obtained from two hundred seventy comprehensive test scores. The scores were reduced to point averages for more ready comparisons.[1]

TABLE III.—POINT AVERAGES IN PSYCHOLOGY OF EXPERIMENTAL AND STANDARD GROUPS

$N = 50$	Point averages and standard deviations at end of marking periods									
	Five weeks		Eighteen weeks		Diff. $m_2 - m_1$ / PE diff.	Significance	Thirty-six weeks		Diff. $m_3 - m_1$ / PE diff.	Significance
	m_1	σ	m_2	σ			m_3	σ		
Experimental.....	2.20	.53	2.47	.56	3.75	99	2.48	.71	3.57	99
Standards $N = 270$..	2.26	.63	2.39	.61	3.61	99	2.41	.60	4.21	100
Diff. $m_E - m_S$ / PE diff.	−1.07		1.38		....	..	.97			
Significance.....	76		83				75			

The above table is interpreted as described under Tables I and II. Both groups show significant gains during the thirty-six weeks. Under the *five weeks* heading, there appears a difference with seventy-six in one hundred chances in favor of the standard group. Under the *thirty-six weeks* heading, this difference has so changed that the seventy-five in one hundred chances of a true difference is now in favor of the experimental group.

[1] Scores were reduced to *T*-scores with mean of fifty and standard deviation of ten. *A*, *B*, *C*, *D*, and *F* denote same range of *T*-scores in all cases.

The correlations between the marks[1] at the different periods for each group in the first experiment are presented for consideration.

TABLE IV.—CORRELATIONS BETWEEN MARKS RECEIVED AT THE DIFFERENT GRADE PERIODS BY EXPERIMENTAL AND CONTROL GROUPS

	Correlations between marks at designated periods		
	$r_{5.18}$	$r_{5.36}$	$r_{18.36}$
Experimental, $N = 50$ other than psychology	.64 ± .06	.59 ± .06	.70 ± .05
Control, $N = 50$ other than psychology	.77 ± .04	.71 ± .05	.81 ± .03
Experimental, $N = 50$ psychology only	.51 ± .07	.48 ± .07	.66 ± .05
Standards, $N = 270$ psychology only	.75 ± .02	.65 ± .02	.80 ± .01

Interpretation of table: Reading from left to right from the designation, "Experimental, other than psychology," the experimental group's average marks in all subjects excepting psychology offer correlations of (1) $r_{5.18} = .64 \pm .06$ or a correlation between the five and eighteen weeks marks of .64 with a probable error of ±.06, (2) $r_{5.36} = .59 \pm .06$ or a correlation between the five and thirty-six weeks marks of .59 with a probable error of ±.06, and (3) $r_{18.36} = .70 \pm .05$ or a correlation between the eighteen and thirty-six weeks marks of .70 with a probable error of ±.05. The rest of the table lends itself to similar interpretation.

The marks of the experimental group afford lower correlations than those of the control group. Adjustments of study method could account for these differences. The correlations of the psychology marks of the experimental group are lower than the correlation of this group's marks in other courses. The gain in both psychology and other than psychology marks was about the same as provided in Tables I and II. Perhaps the pressure to use new methods in psychology brought the use of both effective and ineffective methods while only the effective were used in other courses. The standards in psychology provided correlations similar to those obtained from other-than-psychology marks of the control group.

Second Experiment.—The second experimental group turned out to be a selected low group. The advisers had divided the freshmen

[1] The experimental groups' psychology marks were obtainable in *T*-scores. The control groups' psychology marks were available in letter form. The letters do not offer satisfactory data for correlation. The *T*-scores in psychology of the group of two hundred seventy students providing the standards were available and provided data for representative correlations for control groups.

in psychology without the knowledge of instructors or experimenter. This classification made equation of groups impossible. Therefore, results from the test scores of two hundred seventy students are used for comparisons rather than scores from an equated control group. The achievement and intelligence records of these two groups are summarized in Table V. The two hundred seventy students are designated as the standard group and their results as standards.

TABLE V.—MEANS AND STANDARD DEVIATIONS FROM PERCENTILES OF ACHIEVEMENT AND INTELLIGENCE TEST SCORES OF STANDARD AND EXPERIMENTAL GROUPS

	Achievement		Intelligence	
	Mean	σ	Mean	σ
Experimental $N = 50$.....	36	17	32	23
Standards $N = 270$.....	50	23	52	26

In Table VI, the experimental group's improvement in psychology during the how-to-study instruction appears even more definite with the selected low group than with the more normal group of the first experiment. More of the experimental work was required in written form and more use was made of the blackboard than in the first experiment. Each individual was required to write perspectives,

TABLE VI.—POINT AVERAGES IN PSYCHOLOGY OF EXPERIMENTAL AND STANDARD GROUPS

$N = 50$	Point averages and standard deviations at end of marking periods									
	Five weeks		Nine weeks		Diff. $m_2 - m_1$ / PE diff.	Significance	Eighteen weeks		Diff. $m_3 - m_1$ / PE diff.	Significance
	m_1	σ	m_2	σ			m_3	σ		
Experimental.....	1.70	.65	2.10	.52	5.12	100	2.50	.61	9.53	100
Standards $N = 270$..	2.26	.63	2.30	.68	1.08	77	2.39	.61	3.61	99
Diff. $m_2 - m_3$ / PE diff.	−8.36		3.22		....	...	1.75			
Significance.....	100		98				88			

applications of psychological principles, and generalizations as class preparation. Each class was then divided into groups which pre-

pared an organization of the individuals' papers on the blackboard. These small groups, five to ten individuals, were readily supervised by the instructor, since their errors and deficiencies in procedure were made clear by the blackboard record. Although both the standard and experimental groups' gains are significant, a significant difference in favor of the standard group at the end of the five-week period changes by the end of the eighteen-week to a difference which has eighty-eight chances in one hundred of being favorable to the experimental group.

The correlations between the marks at the different periods for each group in the second experiment furnished the following data.

TABLE VII.—CORRELATIONS BETWEEN MARKS RECEIVED AT THE DIFFERENT GRADE PERIODS BY EXPERIMENTAL AND CONTROL GROUPS

	Correlations between marks at designated periods		
	$r_{5.9}$	$r_{5.18}$	$r_{9.18}$
Experimental, $N = 50$ other than psychology	.77 ± .04	.64 ± .06	.74 ± .04
Experimental, $N = 50$ psychology only	.48 ± .07	.54 ± .07	.82 ± .03
Standards, $N = 270$ psychology only	.78 ± .02	.75 ± .02	.84 ± .01

One intimation from these correlations is the large variation in the psychology marks of the members of the experimental group during the early part of the how-to-study instruction. The relatively high correlation between nine and eighteen weeks marks in psychology indicates that adjustments were made before the end of the nine weeks period. The low correlations do not designate improvement but variation in either direction from the original position. This experimental group showed no gain in marks in other than psychology courses (Table VIII). The correlations between their marks are lower than those between the control group's marks. This may indicate continued adjustment or be due to the difficulty of providing reliable marks for a low group. The experimental group's other-than-psychology marks present the lowest correlation afforded between marks from the nine and eighteen weeks periods. Perhaps the emphasis upon improvement in psychology caused some periodic variation in the time given to other subjects. There could have been some attempts to apply the how-to-study knowledge with

enough success to have neutralized any loss due to emphasis of psychology.

In Table VIII, both experimental groups' point averages in other than the psychology course are produced. The improvement shown by experimental Group 1 does not recur for experimental Group 2. Although experimental 1 shows ninety-seven chances in one hundred of a true gain at the end of the thirty-six weeks, experimental 2 shows little or fifty-three chances in one hundred of a gain.

TABLE VIII.—POINT AVERAGES IN OTHER THAN PSYCHOLOGY COURSES OF EXPERIMENTAL GROUP 1 AND EXPERIMENTAL GROUP 2

$N = 50$	Point averages and standard deviations at end of marking periods									
	Five weeks		Eighteen weeks		Diff. $m_2 - m_1$ / PE diff.	Significance	Thirty-six weeks		Diff. $m_3 - m_1$ / PE diff.	Significance
	m_1	σ	m_2	σ			m_3	σ		
Experimental 1.....	2.23	.52	2.47	.62	3.15	98	2.46	.72	2.77	97
Experimental 2.....	1.90	.71	1.95	.83	.384	60	1.91	.94	.09	53

To investigate the failure of the low group to transfer the methods studied in psychology, two groups of forty each were obtained from the original group from which the first experimental group was selected. The first of these groups contained students with intelligence and achievement records below the fortieth percentile; the second of

TABLE IX.—COMPARISONS OF POINT AVERAGES IN OTHER THAN PSYCHOLOGY COURSES OF TWO GROUPS FORMED BY DIVIDING THE FIRST EXPERIMENTAL GROUP ACCORDING TO ACHIEVEMENT AND INTELLIGENCE

$N = 40$	Point averages and standard deviations at end of marking periods									
	Five weeks		Eighteen weeks		Diff. $m_2 - m_1$ / PE diff.	Significance	Thirty-six weeks		Diff. $m_3 - m_1$ / PE diff.	Significance
	m_1	σ	m_2	σ			m_3	σ		
Below 40th percentile	1.89	.81	1.96	.85	.56	64	1.86	.85	−.24	57
Above 60th percentile	2.42	.68	2.73	.70	3.00	98	2.81	.73	3.68	99

those above the sixtieth. Table IX shows the results of a comparison of these groups' point averages in other than the psychology course. These data demonstrate again the failure of the lower group to improve in other than the psychology course. The correlations between the marks at the different periods for the groups below the fortieth and above the sixtieth percentiles are presented to provide for further comparison.

TABLE X.—CORRELATIONS BETWEEN MARKS IN OTHER THAN PSYCHOLOGY COURSES RECEIVED AT THE DIFFERENT GRADE PERIODS BY GROUPS WITH INTELLIGENCE AND ACHIEVEMENT RECORDS ABOVE THE SIXTIETH AND BELOW THE FORTIETH PERCENTILES

$N = 40$	Correlation between marks at designated periods		
	$r_{5.18}$	$r_{5.36}$	$r_{18.36}$
Below 40th percentile	.56 ± .07	.52 ± .08	.82 ± .04
Above 60th percentile	.63 ± .06	.44 ± .09	.82 ± .04

The low group's correlations show less variation than those of the upper group. However, the correlations of the low group designate variation which must have been in equal amounts in both directions to result in no significant gain as reported in Table IX. To investigate these variations, Table XI was provided and some of the individual records examined.

Of the one hundred six students in the experimental group of Table XI, one person in the lowest group received a mark of *A* and five others a mark of *B* in psychology. The data from which this table was made show the *A* student's percentiles in both achievement and intelligence to be less than five. No similar occurrence is found in the one hundred six students in the control group. The six students from the experimental group who received these exceptional marks gained from nine to eleven points each from their five weeks' to eighteen weeks' marks; only three of these were gained in psychology. These ten points represent gains of from thirty-seven to sixty-two per cent in the individuals' point averages.

It can be noted in Table XI that *E* and *F* marks are received in only the lowest of the experimental groups. Control-group students of the same instructor received failing marks in the three groups with achievement and intelligence percentiles up to the sixtieth.

Of the six students in the experimental group with achievement and intelligence percentiles above the seventy-ninth, five received marks of *A* and one a mark of *B*. Ten students with like records in the control group received two *A*'s, four *B*'s, and four *C*'s.

Along with the exceptional individual improvements are found similar deteriorations not disclosed by the tabled data. Certain

TABLE XI.—RELATIONSHIPS BETWEEN INTELLIGENCE AND ACHIEVEMENT ENTRANCE RECORDS AND PSYCHOLOGY MARKS OF STUDENTS WITH AND WITHOUT INSTRUCTION IN HOW TO STUDY*

Percentiles, achievement and intelligence	Group	Psychology marks					
		A	*B*	*C*	*D*	*E*	*F*
0 to 19	Experimental.....	1	5	23	4	1	1
	Control.....	..	..	XX	III	..	II
20 to 39	Experimental.....	1	4	16			
	Control.....	1	IV	XII	V	I	I
40 to 59	Experimental.....	6	7	14	4		
	Control.....	IV	VI	XII	V	..	I
60 to 79	Experimental.....	7	6				
	Control.....	VI	IV	VII	II		
80 to 99	Experimental.....	5	1				
	Control.....	II	IV	IV			

* These are taken from the experimenter's records because the measuring devices upon which marks are based were the same and all marks were given on the same *T*-score basis. Arabic numerals designate experimental or how-to-study group. Roman numerals designate control or group not instructed in how to study.

individuals who improved in psychology showed accompanying losses in other subjects. The advancement in one subject may have been at the expense of time for other preparations. Other individuals show gradual waning in achievement. The how-to-study procedure may cause these students to despair of success. A wide divergence in individual results is made apparent at the same levels of intelligence and achievement.

CONCLUSIONS AND SUGGESTED PROBLEMS

The tabled data show more improvement in the school marks of students given instruction in how to study than in marks of students without this instruction. The instruction given was more suited to the better student. Whether or not the teaching could be adapted to the lower group remains to be determined. The improvement does not continue during the second semester after the instruction in how to study had ceased. However, there is no indication of a total reversal to old habits, since the experimental groups continued to show higher averages at the close of the second semester than did the control groups.

Table XI reveals less divergence of marks of accomplishment in psychology at the upper levels when how-to-study is emphasized than in the same course when how-to-study is not emphasized. This same tendency to exercise one's ability to the utmost causes greater divergence at the lower levels. In fact the predictive value of entrance examinations is weakened at the lower level and strengthened at the upper by the how-to-study course. These facts furnish further reason for offering the customarily eliminated lower third a chance to attempt college work.

If one accepts the conclusion that instruction in how-to-study raises academic attainment, the need of further experimentation is obvious. What units of the how-to-study courses are most helpful? Can one determine the definite study-help needs for the individual student? What remedial instruction is most applicable to each deficiency? At what school level should how-to-study instruction be stressed? Why do individuals with similar achievement and intelligence records react differently to a how-to-study program?

The answers to these problems are practically necessitated by the modern increase of desirable knowledge. While one must recognize a limit of achievement, any suggested possibility of raising this limit is worthy of consideration. How-to-study instruction affords one means of attaining this desirable goal.

THE EDUCABILITY OF TRANSIENTS

C. H. SMELTZER

Temple University

AND

C. R. ADAMS

Director of Personnel, Pennsylvania Emergency Relief Administration

GENERAL NATURE OF THE PROBLEM

In the autumn of 1935, forty-three transient boys were admitted to the Thaddeus Stevens Industrial School in Lancaster, Pennsylvania. This group was to receive an intensive six months' course in vocational education under the general sponsorship of the Pennsylvania Emergency Relief Administration.[1] All possibilities of returning these boys to their homes and families were either exhausted or found to be inadvisable. At first it had been planned to select the most superior of their number for the vocational program, but later developments made it possible for the entire group of forty-three to enter the school.

A battery of psychometric tests was administered to these boys at the time of entering upon the training program, and a similar battery was administered to the twenty-five boys who remained throughout the six-month period. This study will summarize some of the background characteristics of these boys, and the changes which took place during the training period as measured by the psychometric tests.

BACKGROUND CHARACTERISTICS OF THE TRANSIENT BOYS

The age range for this group was from seventeen to twenty-three years, with a median age of 19.5 years. The range of school grades completed was from the third to the twelfth, with a median of 9.2.[2]

Forty-six per cent had a legal residence in some other State than Pennsylvania. The general home conditions could be considered abnormal in every case. Seventy-three per cent came from the type of home often referred to as "broken"; that is, a home in which the

[1] At that time the Pennsylvania Emergency Relief Administration had jurisdiction over all transient camps in the State. The camp project was absorbed by the Works Progress Administration in December, 1935.

[2] This is 1.2 years higher than reported by W. H. Brentlinger for a group of transients in Texas. *Journal of Applied Psychology*, Vol. XX, No. 1, February, 1936, p. 109.

parents are separated, or one or both of the parents are deceased. An analysis of the reasons for leaving home shows that thirty-nine per cent stated conflict with the family; twenty-seven per cent left orphanages, having reached the maximum age; twenty-four per cent left to seek employment; and the remaining boys left to seek adventure. These boys came from families with a median of 4.2 children. This is twice as high as the median of 2.0 children per family in Pennsylvania, as reported in the 1932 census. In twenty-five per cent of the cases, these boys are the oldest in the family. The group as a whole represents the 2.8 child.

The occupational history of these transient boys is not particularly good. In no case did any of the boys hold jobs requiring skill, although fifty-eight per cent of them had held full-time positions. The median number of full-time positions held by the group was 1.3, and the median length of time in these positions was 7.9 months. The median monthly income for those who held full-time positions was fifty-five dollars. On the other hand, the positions of the fathers present a better picture. Forty-two per cent held positions of a semi-skilled nature; thirty-one per cent were classified as skilled; while the remainder occupied positions in business or professions.[1]

EDUCATIONAL RESULTS FOR THE TRANSIENT BOYS

At the time the transients were admitted to the trade school, the following battery of tests was administered to the entire group; (1) Terman Group Test of Mental Ability (Form A); (2) New Stanford Achievement Test—Advanced Examination (Form V); (3) Stenquist Mechanical Aptitude Tests I and II; and (4) The Thurstone Personality Schedule.

These tests were selected so that a measure would be obtained of mental capacity, scholastic achievement, mechanical aptitude, and emotional adjustment. The six-month training program was to be primarily vocational in nature, combined with related academic training.

At the end of the six months of training, other forms[2] of the same tests were administered to the boys who had remained at the school throughout the course—twenty-five in number. This arrangement permits a measure of the progress which was made in the abilities

[1] Brentlinger found essentially the same general conditions in the Texas group.

[2] Since there are not two forms of the Stenquist and Thurstone tests, they were repeated at the end of the training period.

mentioned above. The following data are a survey of the entire group admitted to the school and a comparison between the initial and end test results for the group of twenty-five who remained throughout the training period.

RESULTS ON THE TERMAN TEST

The mean score for the entire group at the initial testing was ninety-eight, which compares favorably with the median of one hundred four for ninth-grade students. (The median school grade completed by the entire transient group was 9.2.) The mean score on the pre-testing for the group which completed the six-month training was one hundred seventeen corresponding to a mental age of fifteen years eight months; and for the group which did not complete the training, the mean was ninety-two, corresponding to a mental age of fourteen years four months. Apparently there was a tendency for the boys of higher intelligence to remain in school, although the statistical difference between the means of these two groups is not a significant one. Table I shows a critical ratio of 1.9 between the means. If a mental age of sixteen is regarded as the intelligence level of normal adults, the group of transient boys which remained in school may be said to possess normal intelligence.

At the end of the six-month training period, the mean score for the group of twenty-five boys on the Terman Test was one hundred thirty-three, which corresponds to a mental age of sixteen years four months. The group thus grew eight months in mental age during the six-month training period.

RESULTS ON THE STANFORD ACHIEVEMENT TEST

The mean for the entire group was ninety-two, equivalent to an educational age of fourteen years four months, which may be compared with the 9th-grade norm of ninety-six points. The group which remained in school had a mean of ninety-seven (educational age of fifteen years two months), while the group not completing the work had a mean of eighty-seven (educational age of thirteen years five months). Although the boys who remained represented a superior group on the Stanford Achievement Test, the critical ratio of 1.8 (see Table I) indicates that the difference between the two is not statistically significant.

The group which remained throughout the six-month training period had a final mean score of one hundred, equivalent to an educa-

tional age of fifteen years eight months. Thus, during the training period the boys made a gain of six months in educational age. It is interesting to note that this magnitude of gain resulted when the major attention was focused upon vocational training rather than upon academic achievement.

TABLE I.—SHOWING THE INITIAL TEST DIFFERENCES BETWEEN THE BOYS WHO REMAINED IN SCHOOL AND THOSE WHO LEFT BEFORE THE END OF THE TRAINING PERIOD

	N	*M*	SD	CR
Terman Test.				
Group completing course	25	116.7	40.6	1.9
Group not completing course	18	91.7	42.8	
Stanford Achievement Test.				
Group completing course	25	96.5	13.6	1.8
Group not completing course	18	86.7	20.2	

RESULTS ON THE STENQUIST TESTS

The highest age group on which there are norms on both Stenquist tests is 15.5 years. When the transient group completing the course (chronological age 19.5 years) is compared with the norms, the median percentile rank is eighty-four on the initial testing for Test I and ninety-seven on the initial testing for Test II. A gain during the training period of five percentile points was made on Test I and a gain of two percentile points was made on Test II. The transient boys were considerably superior to the group on which the tests have been standardized. The eighteen boys who did not complete the work had initial test scores five percentile points lower than the group of twenty-five boys who remained on Test I, and six percentile points lower on Test II.

THURSTONE PERSONALITY SCHEDULE RESULTS

The distribution of scores on the initial administration for the twenty-five who remained showed one boy to be extremely well-adjusted, nine to be well-adjusted, ten to be average or normally-adjusted, two to be poorly-adjusted, and three to need psychiatric advice, according to Thurstone's classification. The final testing showed five to be extremely well-adjusted, nine to be well-adjusted, eight to be average or normally-adjusted, three to be poorly-adjusted,

and none in the lowest group needing psychiatric advice. The adjustment level was raised from an average to a well-adjusted classification. Those boys who did not complete the course received somewhat lower adjustment scores on the initial testing as a whole than the boys who remained at the school.

GENERAL SUMMARY

This group of transient boys represents a fair sampling of the one transient camp from which it was selected. These boys compare favorably with other adolescents in intelligence, in scholastic achievement (when they are compared with boys and girls who have the same amount of scholastic training), in mechanical aptitude, and represent a typical distribution of persons in regard to emotional adjustment.

During the six months' training, the group as a whole made eight months' progress on the Terman Group Test, six months' progress on the New Stanford Achievement Test, gained five and two percentile points, respectively, on the Stenquist Mechanical Aptitude Tests I and II, and raised the Thurstone emotional adjustment score from a rating of C to B.

There was no way in which each boy could be compelled to stay in school. As a result, eighteen of the group dropped out sometime during the period. The boys who left tended to be somewhat inferior to those who remained at the school in each of the abilities or traits measured. However, in no case was the difference statistically significant.

Contrary to many subjective statements regarding the characteristics of transient adolescents who have taken to the "road" during the past few years, many of these boys appear to be typical members of a population with unfortunate situations in home environment. Although this study should not be construed as representing the entire transient group, it is characteristic of one of the Pennsylvania camps.

Such a "floating population" will soon present some serious social and educational problems if society does not immediately take steps to bring about some remedial type of training. These transient boys represent a group of late adolescents who are definitely educable.

RELATION OF "SOCIAL DOMINANCE" TO GENERAL INTELLIGENCE

ARTHUR F. DODGE

Assistant Professor Industrial Education, University of Illinois

A comparison of test scores of three hundred clerical workers and one hundred fifty-four sales people indicating significant differences between these two occupational groups with respect to social dominance as measured by the Bernreuter Personality Inventory, was reported in an earlier issue of this Journal.[1] As pointed out in this report, the dominance-submission score appears to measure some trait or characteristic which experienced[2] salespeople tend to possess to a greater degree than experienced[2] clerical workers.

This finding raises the question as to the extent to which this measure of personality is independent of the factor of general intelligence. The original data[3] for this study of clerical workers and salespeople included scores on two tests which are commonly accepted as measures of general intelligence—the Pressey Senior Classification and the O'Connor Vocabulary. Also the amount of formal education as indicated by grade attainment was known and this was also accepted as a measure of general intelligence—not a highly valid measure to be sure, but certainly general intelligence is one of the factors determining grade attainment. Table I gives the intercorrelations between social dominance scores and these measures of general intelligence for a group of six hundred fifty-one experienced workers including the clerical and sales groups previously mentioned.

The third column of Table I indicates that the correlations between social dominance and the three different measures of intelligence are all low—the highest being +.115. Previous studies confirm this low correlation between social dominance scores and individual measures of intelligence. Perry,[4] using Thurstone's Psychological

[1] Dodge, Arthur F.: "Social Dominance of Clerical Workers and Sales-Persons as Measured by the Bernreuter Personality Inventory." *Journal of Educational Psychology*, Vol. XXVIII, 1937, pp. 71–73.

[2] All members of these groups had a minimum of two years' experience and ninety per cent had three or more years' experience.

[3] Dodge, Arthur F.: *Occupational Ability Patterns.* Contributions to Education, No. 658, Teachers College, Columbia University, 1935.

[4] Perry, R. C.: *A Group Factor Analysis of the Adjustment Questionnaire.* University of Southern California Education Monographs, Series 1933–1934, pages 27 and 29.

Examination as his measure of intelligence, reports a correlation of $+.1029 \pm .0500$ for a group of one hundred eighty-one males and a correlation of $-.0399 \pm .0561$ for a group of one hundred forty-seven females. Finch and Nemzek[1] report a correlation of $-.07 \pm .067$ with intelligence as measured by averaging the results of five group tests.

Thus, the low correlation between social dominance scores and individual measures of intelligence seems well established by this and previous studies but, so far as the author is aware, no attempt has been made to determine the correlation between social dominance scores and the factor "*g*" as determined by three or more measures of intelligence. Let us, therefore, use some of the proposed methods of factor analysis to see if these results also confirm the low correlation between social dominance and general intelligence.

TABLE I.—INTERCORRELATIONS

	(2) O'Connor vocabulary	(3) Formal education grade attainment	(4) Social dominance
1. Pressey senior classification.....	+.682	+.369	+.115
2. O'Connor vocabulary.....	.....	+.336	−.004
3. Formal education grade attainment.....	.....	.....	+.104

Spearman has pointed out that any three measures of intelligence determine a "*g*"—probably not the true "*g*" but an approximation of this factor. Now if we use Spearman's formula $r_{1g} = \sqrt{\frac{r_{12} \cdot r_{13}}{r_{23}}}$, we may determine the correlations between our three measures of intelligence (Pressey Senior Classification, O'Connor Vocabulary, and Grade Attainment) and the "*g*" which is determined by these three measures.

These correlations are shown in the first column of Table II. As we should expect, we find comparatively high correlations (+.865 and +.788) between this "*g*" and the two tests intended to measure general intelligence while a relatively low correlation (+.427) is shown between this "*g*" and grade attainment.

[1] Finch, F. A. and Nemzek, C. L.: "The Relationship of the Bernreuter Personality Inventory to Scholastic Achievement and Intelligence." *School and Society*, Vol. XXXVI, pp. 594–596.

Now let us apply the test of tetrad differences to these four measures—the three of general intelligence and the one of social dominance. These tetrad differences shown in Table III are large, two of them being greater than four times their probable errors. This indicates that the social dominance score is independent of this factor "*g*" or that there is overlapping among the specific factors. But whatever the cause of the high tetrad differences, it indicates that any further analysis of these four measures by Spearman's method would be of doubtful value.

Another method of determining correlations with "*g*" has been

TABLE II.—CORRELATIONS OF TESTS WITH COMMON FACTORS DETERMINED BY THREE DIFFERENT METHODS OF ANALYSIS

	$\frac{g}{\sqrt{\frac{r_{12} \cdot r_{13}}{r_{23}}}}$	*g* Thompson's modified Hotelling	γ_1 Hotelling method
1. Pressey senior classification	.865	.890	.880
2. O'Connor vocabulary	.788	.763	.854
3. Formal education grade attainment	.427	.430	.655
4. Social dominance Bernreuter	....	.093	.172
Percentage of total variance	.388	.391	.490

TABLE III.—TETRAD DIFFERENCES

Tetrads	Difference	PE[1]	4 × PE
$r_{12}r_{34} - r_{13}r_{24}$	.0725	.0177	.0708
$r_{12}r_{34} - r_{14}r_{23}$	.0324	.0179	.0716
$r_{13}r_{24} - r_{14}r_{23}$	.0401	.0095	.0380

[1] Computed by Spearman's formula for the probable error of a single tetrad difference.

If Spearman's formula for the probable error of a whole group of tetrad differences is used, the result is .0152.

suggested by Thomson.[1] This is a modification of Hotelling's method of factor analysis and the results are shown in column two of Table II.

[1] Thomson, Godfrey H.: "Hotelling's method modified to give Spearman's *g*." *Journal of Educational Psychology*, Vol. XXV, pp. 366–374.

The third column in this table is the result of Hotelling's method.[1] This method always takes out a factor (γ_1) which accounts for a considerably larger percentage of the total variance than does the factor "*g*." In this case γ_1 accounts for .49 of the total variance of the four measures. Thomson's modified Hotelling method gives results very similar to Spearman's, as will be noted by observing the close agreement between the figures in the first and second columns. The percentages of the total variance accounted for by these two *g*'s are .388[2] and .391 respectively.

CONCLUSIONS

1. The tetrad differences shown in Table III indicate that the social-dominance score *may* be independent of the factor "*g*" shown in the first column of Table II.

2. A comparison of the correlations shown in columns one and two of Table II indicates that these two "*g*" factors are nearly identical, the social dominance score having a correlation of +.093 with the second "*g*."

3. These methods of factor analysis, therefore, confirm the conclusion that the social dominance score is either independent of or has only a low correlation with general intelligence.

---

[1] Hotelling, Harold: "Analysis of a Complex of Statistical Variables into Principal Components." *Journal of Educational Psychology*, Vol. XXIV, pp. 417–441, 498–520.

[2] This figure is based on the assumption that the correlation between the social dominance score and the "*g*" of the first column is zero.

NOTE ON THE ACCURACY OF JUDGMENT IN PLUIP ABILITY RATINGS

C. C. MOORE

Grover, Colo.

It has been thought that if a teacher had a group of pupils in his classes for a certain length of time, he would be able to place the pupils with a fair degree of accuracy on a scale from high to low according to each pupils native ability. This ranking would be the same as one obtained by placing these pupils on a scale according to the position each person occupied in his relation to each other pupil.

In order to determine with what degree of accuracy this may be accomplished the writer had five teachers place a number of pupils in the relative position which the teacher deemed to be the most appropriate for each pupil as compared with the intelligence of every other member of the group.

The instructions given to the teachers were: (1) Each pupil on this list is to be ranked according to what you consider to be his position within the group according to his intelligence as compared with the intelligence of each of the other members of the group; and (2) you may use any method of determining this rank you wish; provided, (*a*) no pupil is to be ranked who is not in at least one of your classes, and (*b*) you may omit any name in this list if you feel that you will not be able satisfactorily to place the pupil to whom it belongs.

At a later date the Otis Group Intelligence Scale, Advanced Examination, Form A, was administered to every pupil whose name appeared in any teacher's final ranking. The results obtained on the Otis examination were correlated with the relative placements as turned in on these teacher's lists. No teacher had any knowledge of previous tests of mental ability which may have been administered to these pupils and the only grades to which he had access were those which the teacher had obtained through his own efforts in working with the pupils.

It was known by each teacher that the other four teachers were participating in the study, but the group participating were asked not to discuss the ratings given by them with each other or with teachers who were not taking part in the study. The variation of

the date on which each teacher turned in his report would tend to indicate that each had worked independently of the others. The group upon whose judgment this study depends consisted of three men and two women.

The groups to be ranked by the teachers were all the pupils belonging to the senior division of a small high school. Since the number to be rated and the number actually rated are given in the following table, it will not be necessary to mention those numbers here.

TABLE I.—THE ACCURACY OF THE SPAN OF JUDGMENT IN RANKING PUPILS ACCORDING TO THEIR NATIVE ABILITY

Teacher's number	Years he taught pupils	Number ranked	Number on list	Per cent of list ranked	r	PE_r
1	1	84	92	91.3	.59	.05
1	1	42	92	45.7	.63	.06
1	2	69	80	86.0	.60	.05
2	1	53	92	57.6	.58	.06
3	1	76	92	82.6	.67	.04
4	1–3	69	92	75.0	.49	.06
5	1	80	92	87.0	.49	.06

Teacher 1 is the only one who was permitted to rank the pupils more than once. This was done for the purpose of determining what the effect would be if the teacher were allowed to make the group more selective. This teacher was allowed to eliminate exactly one-half of the group, retaining those pupils whose native ability seemed to be most evident to him. Teacher 1 was also permitted to rank all the same group of pupils who remained in the school one year later. This was done for the purpose of determining what influence the factor of longer acquaintance with the pupils might have.

SUMMARY OF FACTS

Teacher 3 obtained the highest r and the lowest probable error of the correlation. The correlations obtained by the ranking of Teachers 4 and 5 were the lowest of the correlations even though Teacher 4 had taught the pupils longer than any other teacher. The total number ranked by each teacher when compared with the number on each list would indicate that all the teachers eliminated most of the pupils whose relative positions seemed questionable to them.

CONCLUSIONS

1. It is evident that these data point out the fact that it is very difficult, if not impossible, for the average teacher to rank a large group of pupils accurately according to the relative intelligence of the group.

2. Two or three years of acquaintance with the pupil and his work as compared with an acquaintance of one year seems to have but little effect upon the accuracy with which a teacher ranks pupils.

3. The greater the amount of selectiveness used in choosing a group to be ranked on intelligence seems to have a slight but almost negligible effect upon the size of the obtained *r*.

NOTE ON RECOGNITION *VERSUS* RECALL AS METHODS OF TESTING SPELLING

EVELYN M. STURDYVIN

University of Colorado

In the beginning the ability to recall the correct spelling of a word from a dictated list differs from the ability to select from several alternative spellings the word that is correctly spelled. But one who dictates lists of words to be spelled sees that the student writing words frequently proceeds from the point of dictation with the recognition method. He writes with two or more different combinations of letters the word which has been dictated and, after visual evaluation of the list of spellings, he chooses the one that looks correct.

Since the student does employ recognition as a means of completing work originating with the recall method, it seemed likely that while all the processes of one might not be involved in the other, one method might approximate closely enough the other so as to permit its use instead of the other for all practical purposes.

Therefore, as a part of a larger project in spelling research, the author wished to determine whether there is a sufficiently high correlation between the ability to recognize the correct spelling of a word when it is presented with several alternative spellings and the ability to recall the correct spelling of a word from a dictated list to permit the use of the recognition test as a valid measure of the ability to recall.

One fifty-word duplicate form of a spelling scale standardized by the author during the year 1935–1936 was given by both the recall and the recognition method September 30, 1936, to approximately nine hundred university students enrolled in the first quarter of Freshman English. The recognition form, composed of tabulated errors according to their frequency of occurrence in the papers used for standardizing the scale, and consisting of four alternative spellings, including the correct one which the student was instructed to underline, immediately followed the dictated list.

The papers of three hundred twenty students selected at random from the entire list of papers were used for the study.

The distribution of scores for both methods is shown on p. 395.

Only two students made perfect scores, and they made a perfect score on both forms of the test.

THREE HUNDRED TWENTY SCORES; STEP-INTERVAL—TWO POINTS

Scores	Recognition *F*'s	Recall *F*'s
50	14	12
48	30	16
46	30	22
44	31	33
42	25	24
40	28	21
38	28	21
36	27	26
34	28	20
32	20	18
30	12	19
28	9	21
26	9	14
24	9	3
22	6	13
20	4	6
18	2	9
16	6	5
14	2	6
12	0	6
10	0	0
8	0	2
6	0	2
4	0	0
2	0	1

RECOGNITION NORMS

		Range
Q_3	43.61	13–50
Median	37.86	
Q_1	32.07	

RECALL NORMS

		Range
Q_3	42.18	2–50
Median	35.15	
Q_1	27.24	

Reliability was determined by obtaining the coefficient of correlation between the odd and even items of each group and by applying Brown's "prophecy" formula which gave a coefficient of 0.888 for the papers of the recognition method and 0.919 for those of the recall method.

A coefficient of correlation of $0.884 \pm .008$ was obtained between the recognition method and the recall method.

Since a coefficient of correlation of 0.884 between tests is considered high, the recognition method is a sufficiently valid measure of the ability to spell by recall to replace the latter as a method of testing spelling. The recognition method greatly increases objectivity of testing in spelling, since the words are presented in exactly the same manner to all students. That is not possible by using the recall method even when the same lists of words are given, for there is naturally a great difference in tone quality and the rate of dictation of persons giving the words as well as actual differences in pronunciation. The recognition method saves the time of dictating, the time of writing, and the time required for looking over each separate letter of a word, all of which represent a saving for the examiner, the examinee, and the scorer.

BOOK REVIEWS

FRANCES M. GUANELLA. *Block-Building Activities of Young Children.* New York: Archives of Psychology, No. 174, 1934, pp. 92.

This study designed "to discover the nature of the development that takes place in the young child's activities with building materials" is divided into two main parts. The first part involves diary and photographic records of the block-building of sixty-six children between thirteen months and seven and one-half years of age, done in natural nursery school situations. The second part is based on the observation of children's block-building in carefully controlled laboratory situations.

The most elaborate analysis is given to the first part of the study which leaves much to be desired from the point of view of scientific controls. It cannot be determined from the published report just what data the author was able to cull from the diary records already available in the schools for each of the subjects, or what "important activities of each day for every child" were included in the records. The time available for block-building varied in the different nursery groups from forty to two hundred seventy-six minutes per week; the numbers of each type of block available in each group are not given; nor is there any information as to the number of children using a given supply of blocks simultaneously. Although boys are generally considered to excel girls in interest and facility in block-building, no analysis according to sex is attempted except for sex differences in interest in building railroads. A general analysis of block-building behavior according to the capacities supposedly involved in such activity was made consisting of a list of overlapping categories which might have been drawn up entirely independent of the data.

An interesting and suggestive analysis is given in Chapter II of the genetic development of the methods used in block-building by twelve children between fifty-six and ninety-four weeks of age. It appeared that "linear uses preceded areal, areal preceded three-dimensional. Vertical uses preceeded horizontal. . . . Enclosed space clearly followed linear use, but was not definitely separated in time from the solid areal use." The older children studied built so much more elaborately and the types of elaboration were so diverse that the author was unable, because of the limitations of her data, to recognize stages as she did in the data on the twelve younger subjects.

Quantitative treatment is limited almost entirely to the determination of the median age of appearance of various stages in block-

building. The author points out that "median age of first appearance of each activity cannot be considered a norm since the data used in the study were collected on a highly selected group of children. There is some reason to suppose, however, that the *order of development* of the various activities may be approximately the same for an unselected group." Considerable discussion is devoted to the principles of art and design in children's block constructions as well as to symbolism and to the naming of structures.

In Chapter III very brief treatment is given to four experimental set-ups which might well have been elaborated to constitute the entire study. Although forty-eight children participated in this series of experiments only six of them were used in three of the four experiments. The subjects were observed individually while building in a laboratory room with a constant supply of materials in the presence of the experimenter and a stenographer. Records of time spent in building in this situation showed no trend with age, probably because termination of the experiment was suggested to the children by the examiner after seven minutes. The kinds of blocks used and how they were used as well as language during building and the naming of structures were recorded and each child's product was photographed. The third experiment, described in a page and a half, is suggestive of the possibility of devising a scale to measure accuracy and regularity in placing blocks. The last experiment dealt with seven form boards, two of which were too difficult for the subjects. The appendix contains an excellent collection of eighty-six photographs of children's block constructions.

In evaluating the study as a whole, it may be said that had less dependence been placed on diary records made available through nursery school files, and had more time been devoted to the collection of adequate data in the suggestive laboratory set-ups, especially to the standardization of the scale for measuring constructive ability as in the third experiment, the study might have been a much more worth-while contribution to the scientific literature.

DOROTHEA MCCARTHY.

Fordham University.

MARY COLLINS AND JAMES DREVER. *Psychology and Practical Life.* London: University of London Press, 1936, pp. VII + 307.

This book aims (1) to give the layman a general account of modern applied psychology and (2) to be an elementary textbook in this field.

Individual differences, childhood and adolescence, and measurement of intelligence are covered in addition to the more obvious materials on personality, vocational and advertising psychology, etc. The work is a simplified treatment of the fields, obviously written down for the non-professional reader. It is perhaps generalized to a fault, since few facts are cited to support the statements of trends, principles, and the like. Even the lay reader will be disappointed in the practice of devoting the major part of the text to describing measuring devices and procedures and the almost total failure to consider the results obtained from application of the methods.

The American reader will note especially the frequent failure of the authors to make use of recent well-known literature in various fields. Examples of this are: (1) Citing four per cent for red-green color blind instead of the eight per cent found in a number of recent studies; (2) making little or no use of the large mass of literature on child development coming from institutes of child welfare in America; (3) entirely neglecting the important contribution to vocational psychology from measurement of vocational interests. All will approve, however, of the chapter on work and fatigue. In this the authors draw heavily from the noteworthy reports of the Industrial Fatigue Research Board.

Although the book has serious shortcomings as a text, it can be recommended to the layman as a readable, non-technical statement of trends and principles in applied psychology

MILES A. TINKER.
University of Minnesota.

RICHARD T. LA PIERE AND PAUL R. FARNSWORTH. *Social Psychology.* New York: McGraw-Hill Book Company, 1936, pp. XII + 504.

The authors present a brief historical summary of philosophical concepts pertinent to a consideration of social psychology. They then sketch the psycho-biological and sociological bases upon which they believe social psychology must rest. The development of personality differences and deviations through socialization is discussed at length, eventually leading to a consideration of the situational approach to the study of social behavior.

In this volume, a sociologist and a psychologist have attempted to "consolidate the socio-psychological findings both of psychology and of sociology" and to synthesize these approaches. In this, a majority of readers will doubtless agree, they have succeeded.

Each of the twenty chapters is copiously annotated with reference citations and comments supplementing or augmenting points emphasized in the text. These are appended to each chapter and are sufficiently connected in thought to provide running commentaries on the topics under consideration. Reference citations constitute well-selected bibliographies on numerous topics. Broad familiarity with related experimental literature is shown. A wealth of information, carefully abstracted and tersely reported, is contained in the appendices.

An excellent book—one which should place the study of social psychology on a plane of scientific accuracy not ordinarily attained.

GLEN U. CLEETON.

Carnegie Institute of Technology.

NOV 5 1937